

PA 1022787

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

June 11, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/394,576

FILING DATE: July 08, 2002

REC'D 11 AUG 2003

WIPO

PCT

PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS



N. WOODSON

Certifying Officer

BEST AVAILABLE COPY

07/08/02
U.S. PTO
1051

07-10-682576-076800 A/
A/prev

Attorney Docket No. 54720-8022

Transmittal of Provisional Patent Application for Filing
Certification under 37 CFR §1.10 (if applicable)

EV 057 395 186 US
Express Mail Label Number

July 8, 2002
Date of Deposit

I hereby certify that this Request, enclosed application and any other documents referred to as enclosed herein, are being deposited in an envelope with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date above and addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Lynnea B. Anderson
(Print Name of Person Mailing Application)

Lynnea B. Anderson
(Signature of Person Mailing Application)

**Request for Filing a Provisional Patent
Application Under 37 CFR §1.53(c)**

Box Provisional Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

1101 U.S. PTO
1160/394516
07/08/02

Sir

This is a request for filing a Provisional Application for Patent under 37 CFR §1.53(c).

1. Title of Invention:

SECRETED PROTEINS

2. Inventor Applicant(s):

Last Name	First Name	MI	Residence Address	Citizenship
Pardo	Bruno		Geneva, Switzerland	Italy
Jeandenans	Catherine		Aubonne, Switzerland	France
Bougueret	Lydie		Geneva, Switzerland	France

3. Correspondence Address:

Perkins Coie LLP
P.O. Box 2168
Menlo Park, CA 94026
Telephone (650) 838-4300
Customer No. 22918

4. Enclosed documents accompanying this transmittal sheet:

- | | |
|--|---|
| <input checked="" type="checkbox"/> Return Post card for USPTO Date Stamp | <input checked="" type="checkbox"/> Cert. of Express Mail |
| <input checked="" type="checkbox"/> Specification Number of pages <u>83</u> | <input checked="" type="checkbox"/> Drawings - 3 Figures |
| <input checked="" type="checkbox"/> Claims Number of pages <u>03</u> | |
| <input checked="" type="checkbox"/> Sequence Listing Number of pages <u>34</u> | |

PROVISIONAL APPLICATION FILING ONLY

Page 1 of 2

60334576 070002

Attorney Docket No. 54720-8022

5. Government Interest

- This invention was made by an agency of the United States Government, or under contract with an agency of the United States Government. The name of the U.S. Government Agency and the Government contract number are:

6. Method of Payment:

- Applicant claims small entity status. See 37 CFR §1.27
- Please charge Deposit Account No. 50-2207 for the filing fee of \$160. The Commissioner is hereby authorized to charge any deficiency in fees to Deposit Account 50-2207.

Respectfully submitted,

Jacqueline F. Mahoney

Jacqueline F. Mahoney
Registration No. 48,390

July 8, 2002

Correspondence Address:

Customer No. 22918

Phone: 650 838-4300

PROVISIONAL APPLICATION FILING ONLY

Page 2 of 2

PROVISIONAL

CERTIFICATION OF MAILING BY "EXPRESS MAIL"

Express Mail Label: EV057395186US

Date of Deposit: July 8, 2002

Jennifer Anderson

Atty Docket No. 5A720-8022

I hereby certify that this paper or fee and accompanying documents referred to below are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C F R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner of Patents, Washington, D.C. 20231.

5

Secreted ProteinsFIELD OF THE INVENTION

10 The invention relates to secreted proteins and their precursors, isolated polynucleotides encoding such polypeptides, polymorphic variants thereof, and the use of said nucleic acids and proteins or compositions thereof in detection assays, for disease diagnosis, for disease treatment and for drug development.

BACKGROUND

15 Proteins from the phosphatidylethanolamine-binding protein (PEBP) family are highly conserved throughout nature and have no significant sequence homology with other proteins of known structure or function. These proteins are found in a wide variety of neural, peripheral, reproductive and other tissues (Banfield MJ, Barker JJ, Perry AC, Brady RL, Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a 20 role in membrane signal transduction, *Structure*, (1998), Oct 15;6(10):1245-54). They are reported as cytoplasmic proteins, and have been observed as localised to the inner periplasmic membrane (Banfield *et al.*, *supra*)

From a structural point of view, the PEBPs comprise a central beta sheet with highly conserved residues, which contains a binding pocket where a variety of ligands, such as 25 phospholipids, opioids, hydrophobic odorant molecules and possibly G protein coupled signalling proteins can bind. Despite the widespread expression and the resolution of 3D structures of bovine and human PEBP by X-ray crystallography (Scerri L., Vallee B., Burcaud N., Schoentgen F. and Zelwer C., Crystal structure of the phosphatidylethanolamine-binding protein from bovine brain: a novel structural class of phospholipid-binding proteins, *Structure* 1998 Oct 15;6(10):1255-65 and 30 Banfield *et al.*, *supra*), very little is known about the function or properties of the proteins in this family. A possible role for PEBP could be the transduction of membrane signal by regulating intracellular signaling cascades (MAP kinase pathway) (Simister PC, Banfield MJ, Brady RL, The crystal structure of PEBP-2, a homologue of the PEBP/RKIP family, *Acta Crystallogr D Biol*

Crystallogr. (2002) Jun;58(Pt 6 Pt 2):1077-80. A method of inhibiting a protease by contacting it with an effective amount of a PEBP family member has been described (PCT patent publication WO 02/18623).

The present invention discloses a new sub-family of proteins which:

- 5 • are secreted
- have a high level of homology with known PEBP proteins over the ligand binding site region
- have a very different, but conserved, N-ter sequence, which contains the signal peptide for secretion.

10

SUMMARY OF THE INVENTION

The present invention is directed to compositions related secreted proteins, designated 15 herein "Pro412-related" proteins. Such compositions include Pro412-related proteins, having an amino acid sequence of SEQ ID NO: 1 to 4, Pro412-related proteins precursors, Pro412-related antibodies, including monoclonal antibodies and other binding compositions derived therefrom, and methods of making and using these compositions.

In another aspect, the invention includes polynucleotides coding for a polypeptide having 20 an amino acid sequence selected from the group consisting of SEQ ID NO: 1 through 4, antisense oligonucleotides complementary to such sequences, oligonucleotides complementary to Pro412-related gene sequences useful in diagnostic and analytical assays, such as primers for polymerase chain reactions (PCRs), and vectors for expressing Pro412-related proteins.

The invention further includes methods of using Pro412-related proteins compositions, 25 including antisense and antibody compounds, to treat disorders associated with aberrant expression of the Pro412-related proteins of SEQ ID NO: 1 to 4 in an individual, and methods of using Pro412-related proteins compositions, including primers complementary to the Pro412-related gene and/or messenger RNA and anti-Pro412-related antibodies, for detecting and measuring quantities of the Pro412-related proteins in tissues and biological fluids. The invention also 30 further includes methods of screening for compounds that inhibit or increase the expression of Pro412-related proteins, as well as methods of screening for compounds that interact with and/or inhibit or increase the activity of Pro412-related proteins. The invention further encompasses compounds thus identified, as well as compositions thereof.

35 In another aspect, the invention includes polypeptides having a sequence which is at least 95 percent identical to SEQ ID NO: 1 to 4. Preferably, the invention includes polypeptides having at least 97 percent, and more preferably at least 98 percent, and still more preferably at least

99 percent, identity with any one of the sequences selected from SEQ ID NO: 1 to 4. Most preferably, the invention includes polypeptides having a sequence at least 99 percent identical to SEQ ID NO: 1 to 4.

5 In another aspect, the invention includes an isolated peptide fragment of Pro412-related proteins having a sequence identical to a subsequence of SEQ ID NO: 1 to 4.

In another aspect, the invention includes an isolated peptide consisting of 6 to 40 amino acids whose sequence is identical to a subsequence of consecutive amino acids in the Pro412-related protein having the sequence of SEQ ID NO: 1 to 4. Such peptides are useful intermediates
10 in the production of antigenic compositions used in the production of peptide antibodies specific for Pro412-related proteins.

In another aspect, the invention includes isolated antibodies specific for any of the polypeptides, peptide fragments, or peptides described above. Preferably, the antibodies of the invention are monoclonal antibodies. Such antibodies have diagnostic and therapeutic
15 applications, particularly in treating Pro412-related disorders. Treatment methods include, but are not limited to, those that employ antibodies or antibody-derived compositions specific for a Pro412-related antigen. Diagnostic methods for detecting Pro412-related proteins in specific tissue samples, and for detecting levels of expression of Pro412-related proteins in tissues, also form part of the invention. Compositions comprising one or more antibodies described above, together with
20 a pharmaceutically acceptable carrier are also within the scope of the invention.

In another aspect, the invention includes an isolated polynucleotide having a sequence at least 95 percent identical to any sequence of 100 to 1000 consecutive nucleotides of a sequence that encodes a Pro412-related protein having the amino acid sequence of SEQ ID NO: 1 through 4,
25 or any complementary sequence of the foregoing. More preferably, the invention includes an isolated polynucleotide having at least 98 percent, and most preferably at least 99 percent, identity with any sequence of 100 to 1000 consecutive nucleotides of a sequence that encodes a Pro412-related protein having the amino acid sequence of SEQ ID NO: 1 through 4, or any complementary sequence of the foregoing.

30 In another aspect, the invention includes an isolated polynucleotide that forms a detectable duplex under stringent hybridization conditions with a polynucleotide having any sequence of 100 to 1000 consecutive nucleotides of a sequence that encodes a Pro412-related protein having the amino acid sequence of SEQ ID NO: 1 through 4, or any complementary sequence of the foregoing.

35 In another aspect, the invention includes an isolated polynucleotide that encodes Pro412-related proteins having the amino acid sequence of SEQ ID NO: 1 to 4.

In another aspect, the invention includes primer pairs for carrying out a PCR to amplify a

segment of a polynucleotide of the invention. Each primer of a pair is an oligonucleotide having a length of between 15 and 30 nucleotides such that i) one primer of the pair forms a perfectly matched duplex with one strand of a polynucleotide of the invention and the other primer of the pair forms a perfectly matched duplex with the complementary strand of the same polynucleotide, 5 and ii) the primers of a pair form such perfectly matched duplexes at sites on the polynucleotide that are separated by a distance of between 10 and 2500 nucleotides. Preferably, the annealing temperature of each primer of a pair with its respective complementary sequence is substantially the same.

In another aspect, the invention includes natural variants of the Pro412-related protein 10 having a frequency in a selected population of at least two percent. More preferably, such natural variant has a frequency in a selected population of at least five percent, and still more preferably, at least ten percent. Most preferably, such natural variant has a frequency in a selected population of at least twenty percent. The selected population may be any recognized population of study in the field of population genetics. Preferably, the selected population is Caucasian, Negroid, or 15 Asian. More preferably, the selected population is French, German, English, Spanish, Swiss, Japanese, Chinese, Irish, Korean, Singaporean, Icelandic, North American, Israeli, Arab, Turkish, Greek, Italian, Polish, Pacific Islander, Finnish, Norwegian, Swedish, Estonian, Austrian, or Indian. More preferably, the selected population is Icelandic, Saami, Finnish, French of Caucasian ancestry, Swiss, Singaporean of Chinese ancestry, Korean, Japanese, Quebecian, North American 20 Pima Indians, Pennsylvanian Amish and Amish Mennonite, Newfoundlander, or Polynesian.

In another aspect, the invention provides a vector comprising DNA encoding a Pro412-related protein. The invention also includes host cells and transgenic non human animals comprising such a vector. There is also provided a method of making a Pro412-related proteins or 25 Pro412-related proteins precursor. The method comprises the steps of (a) providing a host cell containing an expression vector as disclosed above; (b) culturing the host cell under conditions whereby the DNA segment is expressed; and (c) recovering the protein encoded by the DNA segment. Within one embodiment the expression vector further comprises a secretory signal sequence operably linked to the DNA segment, the cell secretes the protein into a culture medium, and the protein is recovered from the medium.

30

In still a further aspect, the invention includes pharmaceutical compositions and formulations comprising a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1 to 4, and a pharmaceutically acceptable carrier compound.

35

Further aspects of the invention are also described in the specification and in the claims.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

SEQID	content
1	entire protein Q96S96 ('erase' form)
2	mature peptide
3	entire protein Q8WW74 ('grase' form)
4	mature peptide
5	genomic seq.1(NT 008300.6 fragment)
6	genomic seq.2(NT 008300.6 fragment)
7	genomic seq.3(NT 008300.6 fragment)
8	genomic seq.4(NT 008300.6 fragment)
9	genomic seq. for seqID 1 and seqID 2 (AC105046.2 fragment)
10	exon 1
11	exon 2
12	exon 3
13	exon 4
14	exon 5
15	exon 6 for seqID 3 and seqID 4
16	exon 6 for seqID 1 and seqID 2
17	cDNA for seqID 1 and seqID 2 (EST BM984582)
18	cDNA for seqID 3 and seqID 4 (EST AI857902)

BRIEF DESCRIPTION OF THE FIGURES

5 Figure 1 shows the sequence of onc Pro412-related protein of the invention, from SEQ ID NO 1. The tryptic peptides observed by tandem mass spectrometry are bold and underlined. The tryptic peptides observed by MALDI mass spectrometry are in italics and underlined. The signal peptide has been underlined.

10 Figure 2 illustrates the structural homology between the C-term part of the Pro412-related proteins of the invention and the proteins from the PEBP family; pdb1beh designates the human PEBP protein and pdb1a44 designates the bovine PEBP protein. The structure of both proteins has been solved by X-ray crystallography (Banfield *et al.*, *supra* and Serre *et al.*, *supra*). Q96S96 designates one Pro412-related protein of the invention, from SEQ ID NO 2. The secondary structural elements which are predicted to be conserved between the known PEBP proteins and the 15 Pro412-related protein of the invention have been identified above the corresponding stretches of sequence. The cis Glu from the known protein which is mutated to Phe herein has been highlighted.

20 Figure 3 illustrates mainly the conserved N-term part of the Pro412-related proteins of the invention among species. Q96S96 and Q8WW74 correspond to two Pro412-related proteins of the invention, from SEQ ID NO 2 and 4, respectively, and Q9D9G2 and Q9D9L9 are two mouse homologs of the Pro412-related proteins of the invention. The 4 conserved Cys residues have been highlighted, as well as the residue Phe in human and Tyr in mouse which correspond, within the

PEBP-like active site, to the Glu83 residue which is found to be in a *Cis* conformation in human PEBP.

DETAILED DESCRIPTION OF THE INVENTION

5 *Definitions*

As used herein, the term "nucleic acids" and "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. Throughout the present specification, the expression "nucleotide sequence" may be employed to designate indifferently a polynucleotide or a nucleic acid. More precisely, the expression "nucleotide sequence" encompasses the nucleic material itself and is thus not restricted to the sequence information (i.e., the succession of letters chosen among the four base letters) that biochemically characterizes a specific DNA or RNA molecule. Also, used interchangeably herein are terms "nucleic acids", "oligonucleotides", and "polynucleotides".

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated Pro412-related nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. A nucleic acid molecule of the present invention, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of a nucleic acid of the present invention, as a hybridization probe, Pro412-related nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning. A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

35 A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector

and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to Pro412-related nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

5 As used herein, the term "hybridizes to" is intended to describe conditions for moderate stringency or high stringency hybridization, preferably where the hybridization and washing conditions permit nucleotide sequences at least 60% homologous to each other to remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85%, 90%, 95% or 98%
10 homologous to each other typically remain hybridized to each other. Stringent conditions are known to those skilled in the art and can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are as follows: the hybridization step is realized at 65°C in the presence of 6 x SSC buffer, 5 x Denhardt's solution, 0.5% SDS and 100µg/ml of salmon sperm DNA. The
15 hybridization step is followed by four washing steps:
- two washings during 5 min, preferably at 65°C in a 2 x SSC and 0.1%SDS buffer;
- one washing during 30 min, preferably at 65°C in a 2 x SSC and 0.1% SDS buffer,
- one washing during 10 min, preferably at 65°C in a 0.1 x SSC and 0.1%SDS buffer,
these hybridization conditions being suitable for a nucleic acid molecule of about 20 nucleotides in
20 length. It will be appreciated that the hybridization conditions described above are to be adapted according to the length of the desired nucleic acid, following techniques well known to the one skilled in the art, for example be adapted according to the teachings disclosed in Hames B.D. and Higgins S.J. (1985) Nucleic Acid Hybridization: A Practical Approach. Hames and Higgins Ed., IRL Press, Oxford; and Current Protocols in Molecular Bioloy (supra)

25 To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, 90% or 95% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is
30 occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent
35

homology between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % homology = # of identical positions/total # of positions 100).

- The comparison of sequences and determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264-68, modified as in Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-77, the disclosures of which are incorporated herein by reference in their entireties. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J. Mol. Biol. 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to Pro412-related nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to Pro412-related protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Research 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>, the disclosures of which are incorporated herein by reference in their entireties. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989), the disclosures of which are incorporated herein by reference in their entireties. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

- The term "polypeptide" refers to a polymer of amino acids without regard to the length of the polymer; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not specify or exclude post-expression modifications of polypeptides, for example, polypeptides which include the covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like are expressly encompassed by the term polypeptide. Also included within the definition are polypeptides which contain one or more analogs of an amino acid (including, for example, non-naturally occurring amino acids, amino acids which only occur naturally in an unrelated biological system, modified amino acids from mammalian systems etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

The term "protein" as used herein may be used synonymously with the term "polypeptide".

or may refer to, in addition, a complex of two or more polypeptides which may be linked by bonds other than peptide bonds, for example, such polypeptides making up the protein may be linked by disulfide bonds. The term "protein" may also comprehend a family of polypeptides having identical amino acid sequences but different post-translational modifications, such as

- 5 phosphorylations, acylations, glycosylations, and the like, particularly as may be added when such proteins are expressed in eukaryotic hosts.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which
 10 the Pro412-related protein, or a biologically active fragment or homologue thereof is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of a protein according to the invention (e.g. Pro412-related protein, or a biologically active fragment or homologue thereof) in which the protein is separated from cellular components of the cells from which it is isolated or
 15 recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of a protein according to the invention having less than about 30% (by dry weight) of protein other than the Pro412-related protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of protein other than the protein according to the invention, still more preferably less than about 10% of protein other than the
 20 protein according to the invention, and most preferably less than about 5% of protein other than the protein according to the invention. When the protein according to the invention or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

- 25
- The language "substantially free of chemical precursors or other chemicals" includes preparations of Pro412-related protein, or a biologically active fragment or homologue thereof in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical
 30 precursors or other chemicals" includes preparations of a Pro412-related protein having less than about 30% (by dry weight) of chemical precursors or non-Pro412-related proteins chemicals, more preferably less than about 20% chemical precursors or non-Pro412-related proteins chemicals, still more preferably less than about 10% chemical precursors or non-Pro412-related proteins chemicals, and most preferably less than about 5% chemical precursors or non-Pro412-related
 35 proteins chemicals.

The term "recombinant polypeptide" is used herein to refer to polypeptides that have been

artificially designed and which comprise at least two polypeptide sequences that are not found as contiguous polypeptide sequences in their initial natural environment, or to refer to polypeptides which have been expressed from a recombinant polynucleotide.

5 Another aspect of the invention pertains to anti-Pro412-related antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as a Pro412-related protein, or a biologically active fragment or homologue thereof. Examples of immunologically active portions
 10 of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind a Pro412-related protein, or a biologically active fragment or homologue thereof. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an
 15 antigen binding site capable of immunoreacting with a particular epitope of a Pro412-related protein. A monoclonal antibody composition thus typically displays a single binding affinity for a particular Pro412-related protein with which it immunoreacts.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In
 20 the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

35 *Pro412-related proteins of the invention*

The Pro412-related proteins of the invention are listed in SEQ ID NOS 1 to 4, SEQ ID NOS 2 and 4 being the sequences of the mature peptides obtained from SEQ ID NOS 1 and 3,

respectively. The sequences from SEQ ID NOS 1 and 3 represent two highly similar allelic variants, differing only at their C-terminus by the result of one regular Single Nucleotide Polymorphism (SNP) which mutates an E residue to a G residue, and one insertion SNP which induces a frameshift.

- 5 As described in Example 2, the Pro412-related proteins of the invention display in their C-term part a sequence predicted to be folded into an active site like the canonical ones found in the PEBP proteins. This structure has been described to form a ligand binding pocket at one end of the central beta sheet (Banfield *et al., supra*). Interestingly, within this ligand-binding site, the Glu83 in the human PEBP protein (highlighted in Figure 2), which has been shown to adopt a relatively
10 rare *Cis* peptide bond conformation, is mutated to Phe in the Pro412-related proteins of the invention, which is a residue not commonly found in a *Cis* conformation. In the mouse homolog of the Pro412-related proteins of the invention, this residue is mutated to Tyr (see Figure 3).

- Furthermore, as described in Example 3, the N-ter part of the Pro412-related proteins of the invention is found to correspond to a conserved motif, which comprises a very clear signal
15 peptide for secretion of the protein and a conserved repetition of 4 Cys residues.

- The Pro412-related proteins of the invention therefore represent a novel class of proteins which show an active ligand-binding site closely related to the previously characterized PEBP binding site, but which diverge from the PEBP proteins in their N-term, by, firstly, being secreted proteins and not cytoplasmic proteins, and secondly, display a new structural domain conserved
20 amongst many proteins and characterised by a conserved repetition of 4 Cys residues.

Pro412-related nucleic acids

- The human Pro412-related gene sequence, shown herein to comprise 6 exons in SEQ ID
25 NOS 10 to 16, which encodes the Pro412-related proteins. The human Pro412-related gene is localized at chromosome 19. One aspect of the invention pertains to purified or isolated nucleic acid molecules that encode Pro412-related proteins or biologically active portions thereof as further described herein, as well as nucleic acid fragments thereof. Said nucleic acids may be used for example in therapeutic and diagnostic methods and in drug screening assays as further
30 described herein.

- An object of the invention is a purified, isolated, or recombinant nucleic acid coding for a Pro412-related protein, complementary sequences thereto, and fragments thereof. The invention also pertains to a purified or isolated nucleic acid comprising a polynucleotide having at least 95% nucleotide identity with a polynucleotide coding for a Pro412-related protein, advantageously 99
35 % nucleotide identity, preferably 99.5% nucleotide identity and most preferably 99.8% nucleotide identity with a polynucleotide coding for a Pro412-related protein, or a sequence complementary thereto or a biologically active fragment thereof. Another object of the invention relates to

purified, isolated or recombinant nucleic acids comprising a polynucleotide that hybridizes, under the stringent hybridization conditions defined herein, with a polynucleotide coding for a Pro412-related protein, or a sequence complementary thereto or a variant thereof or a biologically active fragment thereof.

- 5 In another preferred aspect, the invention pertains to purified or isolated nucleic acid molecules that encode a portion or variant of a Pro412-related protein, wherein the portion or variant displays a Pro412-related activity of the invention. Preferably said portion or variant is a portion or variant of a naturally occurring full-length Pro412-related protein.

Another object of the invention is a purified, isolated, or recombinant nucleic acid encoding
 10 a Pro412-related protein comprising, consisting essentially of, or consisting of the amino acid sequence of SEQ ID NOS 1 to 4, or fragments thereof, wherein the isolated nucleic acid molecule encodes one or more motifs selected from the group consisting of a Pro412-related target binding region, a signal-anchor sequence and a substrate specificity determining region. For example, the purified, isolated or recombinant nucleic acid may comprise a genomic DNA or fragment thereof
 15 which encodes the polypeptide of SEQ ID NOS 1 to 4 or a fragment thereof, wherein the isolated nucleic acid molecule encodes one or more motifs selected from the group consisting of a Pro412-related target binding region, a nuclear localisation signal and an interferon gamma homology motif. Any combination of said motifs may also be specified. Preferably said Pro412-related target binding region is a target binding region or a DNA binding region.

20 The nucleotide sequence determined from the cloning of the Pro412-related gene allows for the generation of probes and primers designed for use in identifying and/or cloning other Pro412-related proteins family members (e.g. sharing the novel functional domains), as well as Pro412-related proteins homologues from other species.

A nucleic acid fragment encoding a "biologically active portion of a Pro412-related protein" can be prepared by isolating a portion of a nucleotide sequence coding for a Pro412-related protein, which encodes a polypeptide having a Pro412-related protein's biological activity, expressing the encoded portion of the Pro412-related protein (e.g., by recombinant expression *in vitro* or *in vivo*) and assessing the activity of the encoded portion of the Pro412-related protein.

The invention further encompasses nucleic acid molecules that differ from the Pro412-related nucleotide sequences of the invention due to degeneracy of the genetic code and encode the same Pro412-related proteins and fragment of the invention.

In addition to the Pro412-related nucleotide sequences described above, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the Pro412-related proteins may exist within a population (e.g., the human population). Such genetic polymorphism may exist among individuals within a population due to natural allelic variation. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the Pro412-related gene or nucleic acid sequence encoding Pro412-

related protein.

- Nucleic acid molecules corresponding to natural allelic variants and homologues of the Pro412-related nucleic acids of the invention can be isolated based on their homology to the Pro412-related nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion 5 thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Probes based on the Pro412-related nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a 10 fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a Pro412-related protein, such as by measuring a level of a Pro412-related proteins-encoding nucleic acid in a sample of cells from a subject e.g., detecting Pro412-related proteins mRNA levels or determining whether a genomic 15 Pro412-related gene has been mutated or deleted.

It will be appreciated that the invention comprises polypeptides having an amino acid sequence encoded by any of the polynucleotides of the invention.

Uses of Pro412-related nucleic acids

Polynucleotide sequences (or the complements thereof) which encode Pro412-related 20 proteins have various applications, including uses as hybridization probes, in chromosome and gene mapping, and in the generation of antisense RNA and DNA. In addition, Pro412-related proteins-encoding nucleic acids are useful as targets for pharmaceutical intervention, e.g. for the development of DNA vaccines, and for the preparation of Pro412-related proteins by recombinant 25 techniques, as described herein. The polynucleotides described herein, including sequence variants thereof, can be used in diagnostic assays. Accordingly, diagnostic methods based on detecting the presence of such polynucleotides in body fluids or tissue samples are a feature of the present invention. Examples of nucleic acid based diagnostic assays in accordance with the present invention include, but are not limited to, hybridization assays, e.g., in situ hybridization, and PCR-based assays. Polynucleotides, including extended length polynucleotides, sequence variants and 30 fragments thereof, as described herein, may be used to generate hybridization probes or PCR primers for use in such assays. Such probes and primers will be capable of detecting polynucleotide sequences, including genomic sequences that are similar, or complementary to, the Pro412-related proteins polynucleotides described herein.

The invention includes primer pairs for carrying out a PCR to amplify a segment of a 35 polynucleotide of the invention. Each primer of a pair is an oligonucleotide having a length of between 15 and 30 nucleotides such that i) one primer of the pair forms a perfectly matched duplex with one strand of a polynucleotide of the invention and the other primer of the pair form a

perfectly match duplex with the complementary strand of the same polynucleotide, and ii) the primers of a pair form such perfectly matched duplexes at sites on the polynucleotide that separated by a distance of between 10 and 2500 nucleotides. Preferably, the annealing temperature of each primer of a pair to its respective complementary sequence is substantially the same.

- 5 Hybridization probes derived from polynucleotides of the invention can be used, for example, in performing *in situ* hybridization on tissue samples, such as fixed or frozen tissue sections prepared on microscopic slides or suspended cells. Briefly, a labeled DNA or RNA probe is allowed to bind its DNA or RNA target sample in the tissue section on a prepared microscopic, under controlled conditions. Generally, dsDNA probes consisting of the DNA of interest cloned
10 into a plasmid or bacteriophage DNA vector are used for this purpose, although ssDNA or ssRNA probes may also be used. Probes are generally oligonucleotides between about 15 and 40 nucleotides in length. Alternatively, the probes can be polynucleotide probes generated by PCR random priming primer extension or *in vitro* transcription of RNA from plasmids (riboprobes). These latter probes are typically several hundred base pairs in length. The probes can be labeled by
15 any of a number of methods, including fluorescent tags, enzymes or radioactive moieties, according to methods well known in the art. The particular detection method will correspond to the type of label utilized on the probe (e.g., autoradiography, X-ray detection, fluorescent or visual microscopic analysis, as appropriate). The reaction can be further amplified *in situ* using immunocytochemical techniques directed against the label of the detector molecule used, such
20 antibodies directed to a fluorescein moiety present on a fluorescently labeled probe, or against avidin, or marker enzymes (peroxidase, alkaline phosphatase). Specific labeling and *in situ* detection methods can be found, for example, in Howard, G. C., Ed., *Methods in Nonradioactive Detection*, Appleton & Lange, Norwalk, Conn., (1993), herein incorporated by reference. One preferred assay for detecting nucleic acids encoding Pro412-related protein utilizes the subject
25 polynucleotides, or fragments thereof, as primers in a PCR-based assay. According to the assay, nucleic acids present in a test tissue or cell sample are amplified by polymerase chain reaction (PCR) using two primers consisting of at least 15 nucleotides derived from a polynucleotide of the invention, including primers derived from variants and/or extensions of such sequences, as described herein. Amplification products are detected in the sample by a method that is appropriate
30 to the particular label used to label the amplification products, according to methods as described in U.S. Pat. No. 4,683,195. For use in PCR detection methods, such as PCR *in situ* hybridization, PCR primers are selected to be at least 15 nucleotides in length, and preferably between about 15 and 30 nucleotides in length, and are selected from the DNA molecule of interest, according to methods known in the art. It may be desirable to select, for primers, sequences that encompass the
35 longer nucleotide sequences. Preferably, the probes are selected such that the two hybridization sites are separated by between about 10 to 1,000 nucleotides (occasionally up to about 10,000 nucleotides). PCR *in situ* hybridization of tissue sections and/or cell samples provides a highly

sensitive detection method for rare cell types in fixed cell or tissue samples. The PCR in situ hybridization detection method is carried out in accordance with methods that are known in the art, e.g., Nuovo, G. J., *PCR IN SITU HYBRIDIZATION: PROTOCOLS AND APPLICATIONS*, Raven Press, N.Y., 1992; U.S. Pat. No. 5,538,871, both of which are incorporated herein by reference.

- 5 Briefly, a cell sample (tissue on microscopic slide, pelleted cell suspension) is fixed using a common fixative preparation, such as buffered formalin, formaldehyde or the like. Proteinase or detergent treatment is favored following fixation, to increase cell permeability to reagents. The PCR reaction is carried out in situ by polymerase chain reaction (PCR) using two primers. As discussed above, the primers are designed to selectively amplify one or more of the nucleotide sequences described herein. The amplification reaction mixture contains, in addition to the target nucleotide sample and the primers, a thermostable DNA polymerase, such as a polymerase derived from *Thermus aquaticus* (Taq polymerase, U.S. Pat. 4,889,818), and a sufficient quantity of the four standard deoxyribonucleotides (dNTPs), one or more of which may be labeled to facilitate detection. The reaction mixture is subjected to several rounds of thermocycling to produce 10 multiple copies (amplification products) of the target nucleotide sequence. Amplification products are then detected in the sample, for example by detecting radioactively labeled amplification products. Hybridization probes and PCR primers may also be selected from the genomic sequences corresponding to the full-length proteins identified in accordance with the present invention, including promoter, enhancer elements and introns of the gene encoding the naturally occurring 15 polypeptide. Nucleotide sequences encoding a Pro412-related protein can also be used to construct hybridization probes for mapping the gene which encodes that Pro412-related protein and for the genetic analysis of individuals. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as in situ hybridization, linkage analysis against known chromosomal markers, and hybridization screening 20 with libraries. Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the Pro412-related protein cDNA. Computer analysis of the 3' untranslated region is used to rapidly select primers that do not span more than one exon in the genomic DNA, which would complicate the amplification process. Individuals carrying variations of, or mutations in the gene encoding a Pro412-related protein of the present invention may be 25 detected at the DNA level by a variety of techniques. Nucleic acids used for diagnosis may be obtained from a patient's cells, including, for example, tissue biopsy and autopsy material. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR [Saiki, et al. *Nature* 324:163-166 (1986)] prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the 30 present invention can be used to identify and analyze mutations in the gene of the present invention. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified 35

DNA to radiolabeled RNA of the invention or alternatively, radiolabeled antisense DNA sequences of the invention. Sequence changes at specific locations may also be revealed by nuclease protection assays, such RNase and S1 protection or the chemical cleavage method [e.g. Cotton, et al., *Proc. Natl. Acad. Sci. USA* 85:4397-4401 (1985)], or by differences in melting temperatures.

- 5 "Molecular beacons" [Kostrikis L. G. et al., *Science* 279:1228-1229 (1998)], hairpin-shaped, single-stranded synthetic oligonucleotides containing probe sequences which are complementary to the nucleic acid of the present invention, may also be used to detect point mutations or other sequence changes as well as monitor expression levels of Pro412-related proteins.

Antisense Compounds. Oligonucleotides of the invention, including PCR primers and antisense compounds, are synthesized by conventional means on a commercially available automated DNA synthesizer, e.g. an Applied Biosystems (Foster City, CA) model 380B, 392 or 394 DNA/RNA synthesizer, or like instrument. Preferably, phosphoramidite chemistry is employed, e.g. as disclosed in the following references: Beaucage and Iyer, *Tetrahedron*, 48: 2223-2311 (1992); Molko et al, U.S. patent 4,980,460; Koster et al, U.S. patent 4,725,677; Caruthers et al, U.S. patents 4,415,732; 4,458,066; and 4,973,679; and the like. For therapeutic use, nuclease resistant backbones are preferred. Many types of modified oligonucleotides are available that confer nuclease resistance, e.g. phosphorothioate, phosphorodithioate, phosphoramidate, or the like, described in many references, e.g. phosphorothioates: Stec et al, U.S. patent 5,151,510; Hirschbein, U.S. patent 5,166,387; Bergot, U.S. patent 5,183,885; phosphoramidates: Froehler et al, International application PCT/US90/03138; and for a review of additional applicable chemistries: Uhlmann and Peyman (cited above). The length of the antisense oligonucleotides, i.e. such contiguous sequence, is sufficiently large to ensure that specific binding will take place only at the desired target polynucleotide and not at other fortuitous sites. The upper range of the length is determined by several factors, including the inconvenience and expense of synthesizing and purifying oligomers greater than about 30-40 nucleotides in length, the greater tolerance of longer oligonucleotides for mismatches than shorter oligonucleotides, and the like. Preferably, the antisense oligonucleotides of the invention have lengths in the range of about 15 to 40 nucleotides. More preferably, the oligonucleotide moieties have lengths in the range of about 18 to 25 nucleotides.

30

Pro412-related proteins

The terms "Pro412-related protein" and "Pro412-related proteins" is used herein to embrace any and all of the proteins and polypeptides of the present invention. Also forming part of the invention are polypeptides encoded by the polynucleotides of the invention, as well as fusion polypeptides comprising such polypeptides. The invention embodies polypeptides encoded by the nucleic acid sequences of the Pro412-related gene, as well as the Pro412-related proteins

from humans, including isolated or purified Pro412-related proteins consisting of, consisting essentially of, or comprising the sequence of SEQ ID NOS 1 to 4.

The present invention embodies isolated, purified, and recombinant polypeptides comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids, more preferably at least 12, 15, 20, 25, 30, 40, 50, or 75 amino acids of SEQ ID NOS 1 to 4. In preferred embodiments the contiguous stretch of amino acids comprises the site of a mutation or functional mutation, including a deletion, addition, swap or truncation of the amino acids in the Pro412-related protein sequence. The invention also concerns the polypeptide encoded by the Pro412-related nucleotide sequences of the invention, or a complementary sequence thereof or a fragment thereof.

One aspect of the invention pertains to isolated Pro412-related proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-Pro412-related antibodies. In one embodiment, native Pro412-related proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, Pro412-related proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a Pro412-related protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

Typically, biologically active portions comprise a domain or motif with at least one activity of a Pro412-related protein. The present invention also embodies isolated, purified, and recombinant portions or fragments of a Pro412-related protein comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids, more preferably at least 12, 15, 20, 25, 30, 40, 50 or 75 amino acids of SEQ ID NOS 1 to 4. Also encompassed are Pro412-related proteins which comprise between 10 and 20, between 20 and 50, between 30 and 60, between 50 and 99 amino acids of SEQ ID NOS 1 to 4. In other preferred embodiments the contiguous stretch of amino acids comprises the site of a mutation or functional mutation, including a deletion, addition, swap or truncation of the amino acids in the Pro412-related protein sequence.

A biologically active Pro412-related protein may, for example, comprise at least 1, 2, 3, 5, 10, 20 or 30 amino acid changes from the sequence of SEQ ID NOS 1 to 4, or may encode a biologically active Pro412-related protein comprising at least 1%, 2%, 3%, 5%, 8%, 10% or 15% changes in amino acids from the sequence of SEQ ID NOS 1 to 4.

In a preferred embodiment, a Pro412-related protein comprises a Pro412-related target binding region and/or a signal-anchor sequence. The invention also concerns the polypeptide encoded by the Pro412-related nucleotide sequences of the invention, or a complementary sequence thereof or a fragment thereof.

In other embodiments, the Pro412-related protein is substantially homologous to the sequences of SEQ ID NOS 1 to 4, and retains the functional activity of the Pro412-related protein, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described

further herein. Accordingly, in another embodiment, the Pro412-related protein is a protein which comprises an amino acid sequence which shares more than about 60% but less than 100% homology with the amino acid sequence of SEQ ID NOS 1 to 4 and retains the functional activity of the Pro412-related proteins of SEQ ID NOS 1 to 4, respectively. Preferably, the protein is at least about 30%, 5 40%, 50%, 60%, 70%, 80%, 85%, 90%, 92%, 95%, 97%, 98%, 99% or 99.8% homologous to SEQ ID NOS 1 to 4, but is not identical to SEQ ID NOS 1 to 4. Preferably the Pro412-related proteins is less than identical (e.g. 100% identity) to a naturally occurring Pro412-related proteins. Percent homology can be determined as further detailed above.

10 *Assessing polypeptides activity*

It will be appreciated that the invention further provides methods of testing the activity of, or obtaining, functional fragments and variants of Pro412-related proteins and Pro412-related nucleotide sequences involving providing a variant or modified Pro412-related proteins, or Pro412-related nucleic acid and assessing whether a polypeptide encoded thereby displays a Pro412-related activity of the invention. Encompassed is thus a method of assessing the function of a Pro412-related proteins polypeptide comprising : (a) providing a Pro412-related protein, or a biologically active fragment or homologue thereof; and (b) testing said Pro412-related protein, or a biologically active fragment or homologue thereof for a Pro412-related activity, preferably a 15 Pro412-related activity. Any suitable format may be used, including cell free, cell-based and in vivo formats. For example, said assay may comprise expressing a Pro412-related nucleic acid in a host cell, and observing Pro412-related activity in said cell. In another example, a Pro412-related protein, or a biologically active fragment or homologue thereof is introduced to a cell, and a 20 Pro412-related activity is observed.

25 A Pro412-related activity may be any activity as described herein, such as (1) mediating levels of a Pro412-related proteins substrate protein; (2) mediating levels of a Pro412-related proteins substrate cleavage product; (3) mediating cell proliferation or cell cycle; or (4) interaction with a Pro412-related target molecule or preferably Pro412-related domain target molecule, 30 preferably interaction with a protein e.g. a substrate protein). Detecting Pro412-related activity may also comprise detecting any suitable therapeutic endpoint discussed herein in the section titled "Methods of Treatment".

Methods for obtaining variant nucleic acids and polypeptides

35 In addition to naturally-occurring allelic variants of the Pro412-related proteins sequences that may exist in the population, the skilled artisan will appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NOS 1 to 4, thereby leading to changes in the

amino acid sequence of the encoded Pro412-related proteins, with or without altering the functional ability of the Pro412-related proteins.

Several types of variants are contemplated including 1) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue and 5 such substituted amino acid residue may or may not be one encoded by the genetic code, or 2) one in which one or more of the amino acid residues includes a substituent group, or 3) one in which the mutated Pro412-related protein is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or 4) one in which the additional amino acids are fused to the mutated Pro412-related protein, such as a leader or signal-10 anchor sequence or a sequence which is employed for purification of the mutated Pro412-related protein or a preprotein sequence. Such variants are deemed to be within the scope of those skilled in the art.

For example, nucleotide substitutions leading to amino acid substitutions can be made in the sequences that do not substantially change the biological activity of the protein. An amino acid 15 residue can be altered from the wild-type sequence encoding a Pro412-related protein, or a biologically active fragment or homologue thereof without altering the biological activity. In general, amino acid residues that are conserved among the Pro412-related proteins of the present invention, are predicted to be less amenable to alteration.

In one aspect, the invention pertains to nucleic acid molecules encoding Pro412-related 20 proteins, or biologically active fragments or homologues thereof that contain changes in amino acid residues that are not essential for activity. Such Pro412-related proteins differ in amino acid sequence from SEQ ID NOS 1 to 4 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 60% homologous to an amino acid sequence 25 selected from the group consisting of SEQ ID NOS 1 to 4. Preferably, the protein encoded by the nucleic acid molecule is at least about 65-70% homologous to an amino acid sequence selected from the group consisting of SEQ ID NOS 1 to 4, more preferably sharing at least about 75-80% identity with an amino acid sequence selected from the group consisting of SEQ ID NOS 1 to 4, even more preferably sharing at least about 85%, 90%, 92%, 95%, 97%, 98%, 99% or 99.8% 30 identity with an amino acid sequence selected from the group consisting of SEQ ID NOS 1 to 4.

In another aspect, the invention pertains to nucleic acid molecules encoding Pro412-related 35 proteins that contain changes in amino acid residues that result in increased biological activity, or a modified biological activity. In another aspect, the invention pertains to nucleic acid molecules encoding Pro412-related proteins that contain changes in amino acid residues that are essential for a Pro412-related activity. Such Pro412-related proteins differ in amino acid sequence from SEQ ID NOS 1 to 4 and display reduced activity, or essentially lack one or more Pro412-related proteins biological activities.

An isolated nucleic acid molecule encoding a Pro412-related protein, or a biologically active fragment or homologue thereof, homologous to a protein of any one of SEQ ID NOS 1 to 4 can be created by introducing one or more nucleotide substitutions, additions or deletions into the 5 nucleotide sequence of SEQ ID NOS 1 to 4 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into any of SEQ ID NOS 1 to 4, by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. For example, conservative amino acid substitutions may be made at one or 10 more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These 15 families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a Pro412-related protein, or a 20 biologically active fragment or homologue thereof may be replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a Pro412-related protein's coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for Pro412-related proteins' biological activity to identify mutants that retain activity. Following mutagenesis of one of SEQ ID NOS 1 to 4, the encoded protein can be expressed recombinantly and the activity of the protein can 25 be determined.

In a preferred embodiment, a mutant Pro412-related domain polypeptide, or a biologically active fragment or homologue thereof encoded by a Pro412-related nucleic acid of the invention can be assayed for a Pro412-related activity in any suitable assay, examples of which are provided herein.

The invention also provides Pro412-related proteins chimeric or fusion proteins. As used 30 herein, a Pro412-related protein's "chimeric protein" or "fusion protein" comprises a Pro412-related protein of the invention or fragment thereof, operatively linked, preferably fused in frame, to a non-Pro412-related protein. In a preferred embodiment, a Pro412-related fusion protein comprises at least one biologically active portion of a Pro412-related protein. In another preferred embodiment, a Pro412-related fusion protein comprises at least two biologically active portions of a Pro412-related protein. For example, in one embodiment, the fusion protein is a GST-Pro412-related 35 fusion protein in which a Pro412-related domain sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant Pro412-related

proteins. In another embodiment, the fusion protein is a Pro412-related protein containing a heterologous signal sequence at its N-terminus, such as for example to allow for a desired cellular localization in a certain host cell.

- The Pro412-related fusion proteins of the invention can be incorporated into
- 5 pharmaceutical compositions and administered to a subject *in vivo*. Moreover, the Pro412-related fusion proteins of the invention can be used as immunogens to produce anti-Pro412-related antibodies in a subject, to purify Pro412-related proteins ligands and in screening assays to identify molecules which inhibit the interaction of Pro412-related proteins polypeptide with a Pro412-related target molecule.
- 10 Furthermore, isolated peptidyl portions of the subject Pro412-related proteins can also be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase F-Moc or t-Boc chemistry. For example, a Pro412-related protein of the present invention may be arbitrarily divided into
- 15 fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a Pro412-related activity, such as by microinjection assays or *in vitro* protein binding assays. In an illustrative embodiment, peptidyl portions of a Pro412-related
- 20 protein, such as a Pro412-related target binding region, can be tested for Pro412-related activity by expression as thioredoxin fusion proteins, each of which contains a discrete fragment of the Pro412-related protein (see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/02502, the disclosures of which are incorporated herein by reference).

- The present invention also pertains to variants of the Pro412-related proteins which
- 25 function as either Pro412-related proteins mimetics or as Pro412-related inhibitors. Variants of the Pro412-related proteins can be generated by mutagenesis, e.g., discrete point mutation or truncation of a Pro412-related protein. An agonist of a Pro412-related protein can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a Pro412-related protein. An antagonist of a Pro412-related protein can inhibit one or more of the
- 30 activities of the naturally occurring form of the Pro412-related protein by, for example, competitively inhibiting the association of Pro412-related protein with a Pro412-related target molecule. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, variants of a Pro412-related protein which function as either Pro412-related proteins agonists (mimetics) or as Pro412-related proteins antagonists can be identified by
- 35 screening combinatorial libraries of mutants, e.g., truncation mutants, of a Pro412-related protein for Pro412-related protein agonist or antagonist activity. In one embodiment, a variegated library of Pro412-related proteins variants is generated by combinatorial mutagenesis at the nucleic acid

level and is encoded by a variegated gene library. A variegated library of Pro412-related proteins variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential Pro412-related proteins sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of Pro412-related proteins sequences therein. There are a variety of methods which can be used to produce libraries of potential Pro412-related proteins variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential Pro412-related proteins sequences.

In addition, libraries of fragments of a Pro412-related protein coding sequence can be used to generate a variegated population of Pro412-related proteins fragments for screening and subsequent selection of variants of a Pro412-related protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of Pro412-related proteins coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the Pro412-related protein.

Modified Pro412-related proteins can be used for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo). Such modified peptides, when designed to retain at least one activity of the naturally occurring form of the protein, are considered functional equivalents of the Pro412-related protein described in more detail herein. Such modified peptide can be produced, for instance, by amino acid substitution, deletion, or addition.

Whether a change in the amino acid sequence of a peptide results in a functional Pro412-related proteins homolog (e.g. functional in the sense that it acts to mimic or antagonize the wild-type form) can be readily determined by assessing the ability of the variant peptide to cleave one or more substrate proteins, or produce a response in cells, in a fashion similar to the wild-type Pro412-related protein or competitively inhibit such a response. Peptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further contemplates a method of generating sets of combinatorial mutants of the presently disclosed Pro412-related proteins, as well as truncation and fragmentation mutants, and is especially useful for identifying potential variant sequences which are functional in

binding to or cleaving a Pro412-related protein- target protein but differ from a wild-type form of the protein by, for example, efficacy, potency and/or intracellular half-life. One purpose for screening such combinatorial libraries is, for example, to isolate novel Pro412-related proteins homologs which function as either an agonist or an antagonist of the biological activities of the 5 wild-type protein, or alternatively, possess novel activities all together. For example, mutagenesis can give rise to Pro412-related proteins homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. The altered protein can be rendered either more stable or less stable to activation of the protease proenzyme, proteolytic degradation, or cellular process which result in destruction of, or otherwise inactivation of, a Pro412-related protein. Such 10 Pro412-related proteinshomologs, and the genes which encode them, can be utilized to alter the envelope of expression for a particular recombinant Pro412-related protcinsprotein by modulating the half-life of the recombinant protein. For instance, a short half-life can give rise to more transient biological effects associated with a particular recombinant Pro412-related protein and, when part of an inducible expression system, can allow tighter control of recombinant protein 15 levels within a cell. As above, such proteins, and particularly their recombinant nucleic acid constructs, can be used in gene therapy protocols.

In an illustrative embodiment of this method, the amino acid sequences for a population of Pro412-related protcins homologs or other related proteins are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, Pro412-related 20 proteins homologs from one or more species, or Pro412-related protcins homologs from the same species but which differ due to mutation. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. There are many ways by which the library of potential Pro412-related proteins homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be 25 carried out in an automatic DNA synthesizer, and the synthetic genes then be ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential Pro412-related proteins sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) Tetrahedron 393; Itakura et al. (1981) Rcombinant DNA, Proc 3rd Cleveland 30 Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier pp. 273-289; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nuclic Acid Res. 11:477 Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) Science 249:386-390; Roberts et al. (1992) PNAS 89 2429-2433; Devlin et al (1990) Science 249: 404-406; Cwirka et al. (1990) PNAS 87: 6378-6382; as 35 well as U.S Patents Nos: 5, 223,409, 5,198,346, and 5,096,815). The disclosures of the above references are incorporated herein by reference in their entireties.

Alternatively, other forms of mutagenesis can be utilized to generate a combinatorial

library, particularly where no other naturally occurring homologs have yet been sequenced. For example, Pro412-related proteins homologs (both agonist and antagonist forms) can be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al. (1994) Biochemistry 33:1565-1572; Wang et al. (1994) J Biol. Chem. 269:3095-3099; Balint et al. (1993) Gene 137:109-118; Grodberg et al. (1993) Eur. J Biochem. 218:597-601; Nagashima et al. (1993) J Biol. Chem. 268:2888-2892; Lowman et al. (1991) Biochemistry 30:10832-10838; and Cunningham et al. (1989) Science 244:1081-1085), by linker scanning mutagenesis (Gustin et al. (1993) Virology 193:653-660; Brown et al. (1992) Mol. Cell Biol. 12:2644 2652; McKnight et al. (1982) Science 232:316); by saturation mutagenesis (Meyers et al. (1986) Science 232:613); by PCR mutagenesis (Leung et al. (1989) Method Cell Mol Biol 1: 1-19); or by random mutagenesis (Miller et al. (1992) A Short Course in Bacterial Genetics, CSHL Press, Cold Spring Harbor, NY; and Greener et al. (1994) Strategies in Mol Biol 7:32-34, the disclosures of which are incorporated herein by reference in their entireties).

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations, as well as for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of Pro412-related proteins. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected.

Each of the illustrative assays described below are amenable to high throughput analysis as necessary to screen large numbers of degenerate Pro412-related proteins sequences created by combinatorial mutagenesis techniques. In one screening assay, the candidate gene products are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind a Pro412-related target molecule (for example a modified peptide substrate) via this gene product is detected in a "panning assay". For instance, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) BioTechnology 9:1370-1371, and Goward et al. (1992) TIBS 18:136 140). In a similar fashion, fluorescently labeled Pro412-related target can be used to score for potentially functional Pro412-related proteins homologs. Cells can be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence- activated cell sorter.

In an alternate embodiment, the gene library is expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First,

since these phage can be applied to affinity matrices at very high concentrations, a large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli*

5 filamentous phages M13, fd, and fl are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) J Biol. Chem. 267:16007-16010; Griffiths et al. (1993) EMBO J 12:725-734; Clackson et al. (1991) Nature 352:624-628; and Barbas et al.

10 (1992) PNAS 89:4457 4461, the disclosures of which are incorporated herein by reference in their entireties). In an illustrative embodiment, the recombinant phage antibody system (RPAS, Pharmacia Catalog number 27-9400-01) can be easily modified for use in expressing Pro412-related proteins combinatorial libraries, and the Pro412-related proteins phage library can be panned on immobilized Pro412-related target molecule (glutathione immobilized Pro412-related

15 target-GST fusion proteins or immobilized DNA). Successive rounds of phage amplification and panning can greatly enrich for Pro412-related proteins homologs which retain an ability to bind a Pro412-related target and which can subsequently be screened further for biological activities in automated assays, in order to distinguish between agonists and antagonists.

The invention also provides for identification and reduction to functional minimal size of the Pro412-related proteins functional domains, to generate mimetics, e.g. peptide or non-peptide agents, which are able to disrupt binding of a polypeptide of the present invention with a Pro412-related target molecule (protein or DNA). Thus, such mutagenic techniques as described above are also useful to map the determinants of Pro412-related proteins which participate in protein-protein or protein-DNA interactions involved in, for example, binding to a Pro412-related target protein or

25 DNA. To illustrate, the critical residues of a Pro412-related protein which are involved in molecular recognition of the Pro412-related target can be determined and used to generate Pro412-related target-13P-derived peptidomimetics that competitively inhibit binding of the Pro412-related protein to the Pro412-related target. By employing, for example, scanning mutagenesis to map the amino acid residues of a particular Pro412-related protein involved in binding a Pro412-related

30 target, peptidomimetic compounds can be generated which mimic those residues in binding to a Pro412-related target, and which, by inhibiting binding of the Pro412-related protein to the Pro412-related target molecule, can interfere with the function of a Pro412-related protein in transcriptional regulation of one or more genes. For instance, non hydrolyzable peptide analogs of such residues can be generated using retro-inverse peptides (e.g., see U.S. Patents 5,116,947 and

35 5,219,089; and Pallai et al. (1983) Int J Pept Protein Res 21:84-92), benzodiazepine (e.g., see Freidinger et al. in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in Peptides.- Chemistry and Biology,

- G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in Peptides: Chemistry and Biology, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) J Med Chem 29:295; and Ewenson et al. in Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), P-turn dipeptide cores (Nagai et al. (1985) Tetrahedron Lett 26:647; and Sato et al. (1986) J Chem Soc Perkin Trans 1: 123 1), and P-aminoalcohols (Gordon et al. (1985) Biochem Biophys Res Commun 126:419; and Dann et al. (1986) Biochem Biophys Res Commun 134:71, the disclosures of which are incorporated herein by reference in their entireties).
- 10 An isolated Pro412-related protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind Pro412-related proteins using standard techniques for polyclonal and monoclonal antibody preparation. A full-length Pro412-related protein can be used or, alternatively, the invention provides antigenic peptide fragments of Pro412-related proteins for use as immunogens. Any fragment of a Pro412-related protein which contains at least one
- 15 antigenic determinant may be used to generate antibodies. The antigenic peptide of a Pro412-related protein comprises at least 6 amino acid residues of an amino acid sequence selected from the group consisting of SEQ ID NOS 1 to 4, and encompasses an epitope of a Pro412-related protein such that an antibody raised against the peptide forms a specific immune complex with a Pro412-related protein. Preferably, the antigenic peptide comprises at least 8 or 10 amino acid
- 20 residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of a Pro412-related protein that are located on the surface of the protein, e.g., hydrophilic regions.

- 25 A Pro412-related protein immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed Pro412-related protein or a chemically synthesized Pro412-related protein. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar
- 30 immunostimulatory agent. Immunization of a suitable subject with an immunogenic Pro412-related protein preparation induces a polyclonal anti-Pro412-related antibody response.

The invention concerns antibody compositions, either polyclonal or monoclonal, capable of selectively binding, or selectively bind to an epitope-containing a polypeptide comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids. The invention also concerns a purified or isolated antibody capable of specifically binding to a mutated Pro412-related protein or to a fragment or variant thereof comprising an epitope of the mutated Pro412-related protein.

Primers and probes

Primers and probes of the invention can be prepared by any suitable method, including, for example, cloning and restriction of appropriate sequences and direct chemical synthesis by a method such as the phosphodiester method of Narang SA et al (Methods Enzymol 1979;68:90-98), the phosphodiester method of Brown EL et al (Methods Enzymol 1979;68:109-151), the diethylphosphoramidite method of Beaucage et al (Tetrahedron Lett 1981, 22: 1859-1862) and the solid support method described in EP 0 707 592, the disclosures of which are incorporated herein by reference in their entireties.

Detection probes are generally nucleic acid sequences or uncharged nucleic acid analogs such as, for example peptide nucleic acids which are disclosed in International Patent Application WO 92/20702, morpholino analogs which are described in U.S. Patents Numbered 5,185,444; 5,034,506 and 5,142,047. If desired, the probe may be rendered "non-extendable" in that additional dNTPs cannot be added to the probe. In and of themselves analogs usually are non-extendable and nucleic acid probes can be rendered non-extendable by modifying the 3' end of the probe such that the hydroxyl group is no longer capable of participating in elongation. For example, the 3' end of the probe can be functionalized with the capture or detection label to thereby consume or otherwise block the hydroxyl group.

Any of the polynucleotides of the present invention can be labeled, if desired, by incorporating any label known in the art to be detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include radioactive substances (including, ³²P, ³⁵S, ³H, ¹²⁵I), fluorescent dyes (including, 5-bromodesoxyuridin, fluorescein, acetylaminofluorene, digoxigenin) or biotin. Preferably, polynucleotides are labeled at their 3' and 5' ends. Examples of non-radioactive labeling of nucleic acid fragments are described in (Urdea et al. (Nucleic Acids Research. 11:4937-4957, 1988) or Sanchez-Pescador et al. (J Clin Microbiol. 26(10):1934-1938, 1988). In addition, the probes according to the present invention may have structural characteristics such that they allow the signal amplification, such structural characteristics being, for example, branched DNA probes as those described by Urdea et al (Nucleic Acids Symp. Ser. 24:197-200, 1991) or in the European patent No. EP 0 225 807 (Chiron).

A label can also be used to capture the primer, so as to facilitate the immobilization of either the primer or a primer extension product, such as amplified DNA, on a solid support. A capture label is attached to the primers or probes and can be a specific binding member which forms a binding pair with the solid's phase reagent's specific binding member (e.g. biotin and streptavidin). Therefore depending upon the type of label carried by a polynucleotide or a probe, it may be employed to capture or to detect the target DNA. Further, it will be understood that the

polynucleotides, primers or probes provided herein, may, themselves, serve as the capture label. For example, in the case where a solid phase reagent's binding member is a nucleic acid sequence, it may be selected such that it binds a complementary portion of a primer or probe to thereby immobilize the primer or probe to the solid phase. In cases where a polynucleotide probe itself 5 serves as the binding member, those skilled in the art will recognize that the probe will contain a sequence or "tail" that is not complementary to the target. In the case where a polynucleotide primer itself serves as the capture label, at least a portion of the primer will be free to hybridize with a nucleic acid on a solid phase. DNA labeling techniques are well known to the skilled technician.

10 The probes of the present invention are useful for a number of purposes. They can be notably used in Southern hybridization to genomic DNA. The probes can also be used to detect PCR amplification products. They may also be used to detect mismatches in the Pro412-related gene or a Pro412-related protein mRNA using other techniques.

Any of the nucleic acids, polynucleotides, primers and probes of the present invention can 15 be conveniently immobilized on a solid support. Solid supports are known to those skilled in the art and include the walls of wells of a reaction tray, test tubes, polystyrene beads, magnetic beads, nitrocellulose strips, membranes, microparticles such as latex particles, sheep (or other animal) red blood cells, duracytes and others. The solid support is not critical and can be selected by one skilled in the art. Thus, latex particles, microparticles, magnetic or non-magnetic beads, 20 membranes, plastic tubes, walls of microtiter wells, glass or silicon chips, sheep (or other suitable animal's) red blood cells and duracytes are all suitable examples. Suitable methods for immobilizing nucleic acids on solid phases include ionic, hydrophobic, covalent interactions and the like. A solid support, as used herein, refers to any material which is insoluble, or can be made insoluble by a subsequent reaction. The solid support can be chosen for its intrinsic ability to 25 attract and immobilize the capture reagent. Alternatively, the solid phase can retain an additional receptor which has the ability to attract and immobilize the capture reagent. The additional receptor can include a charged substance that is oppositely charged with respect to the capture reagent itself or to a charged substance conjugated to the capture reagent. As yet another alternative, the receptor molecule can be any specific binding member which is immobilized upon 30 (attached to) the solid support and which has the ability to immobilize the capture reagent through a specific binding reaction. The receptor molecule enables the indirect binding of the capture reagent to a solid support material before the performance of the assay or during the performance of the assay. The solid phase thus can be a plastic, derivatized plastic, magnetic or non-magnetic metal, glass or silicon surface of a test tube, microtiter well, sheet, bead, microparticle, chip, sheep 35 (or other suitable animal's) red blood cells, duracytes and other configurations known to those of ordinary skill in the art. The nucleic acids, polynucleotides, primers and probes of the invention can be attached to or immobilized on a solid support individually or in groups of at least 2, 5, 8, 10,

12, 15, 20, or 25 distinct polynucleotides of the invention to a single solid support. In addition, polynucleotides other than those of the invention may be attached to the same solid support as one or more polynucleotides of the invention.

- Any polynucleotide provided herein may be attached in overlapping areas or at random
- 5 locations on a solid support. Alternatively the polynucleotides of the invention may be attached in an ordered array wherein each polynucleotide is attached to a distinct region of the solid support which does not overlap with the attachment site of any other polynucleotide. Preferably, such an ordered array of polynucleotides is designed to be "addressable" where the distinct locations are recorded and can be accessed as part of an assay procedure. Addressable polynucleotide arrays
- 10 typically comprise a plurality of different oligonucleotide probes that are coupled to a surface of a substrate in different known locations. The knowledge of the precise location of each polynucleotides location makes these "addressable" arrays particularly useful in hybridization assays. Any addressable array technology known in the art can be employed with the polynucleotides of the invention. One particular embodiment of these polynucleotide arrays is
- 15 known as the Genechips, and has been generally described in US Patent 5,143,854; PCT publications WO 90/15070 and 92/10092, the disclosures of which are incorporated herein by reference in their entireties.

Chemical Manufacture of Pro412-related Compositions

20

- Peptides of the invention are synthesized by standard techniques, e.g. Stewart and Young, Solid Phase Peptide Synthesis, 2nd Ed. (Pierce Chemical Company, Rockford, IL, 1984). Preferably, a commercial peptide synthesizer is used, e.g. Applied Biosystems, Inc. (Foster City, CA) model 430A, and polypeptides of the invention may be assembled from multiple, separately
- 25 synthesized and purified, peptide in a convergent synthesis approach, e.g. Kent et al, U.S. patent 6,184,344 and Dawson and Kent, Annu. Rev. Biochem., 69: 923-960 (2000). Peptides of the invention may be assembled by solid phase synthesis on a cross-linked polystyrene support starting from the carboxyl terminal residue and adding amino acids in a stepwise fashion until the entire peptide has been formed. The following references are guides to the chemistry employed during
- 30 synthesis: Schnolzer et al, Int. J. Peptide Protein Res., 40: 180-193 (1992); Merrifield, J. Amer. Chem. Soc., Vol. 85, pg. 2149 (1963); Kent et al., pg 185, in Peptides 1984, Ragnarsson, Ed (Almquist and Wiksell, Stockholm, 1984); Kent et al., pg. 217 in Peptide Chemistry 84, Izumiya, Ed. (Protein Research Foundation, B.H. Osaka, 1985); Merrifield, Science, Vol. 232, pgs. 341-347 (1986); Kent, Ann. Rev. Biochem., Vol. 57, pgs. 957-989 (1988), and references cited in these
- 35 latter two references.

Preferably, chemical synthesis of polypeptides of the invention is carried out by the assembly of peptide fragments by native chemical ligation, as described by Dawson et al, Science,

266: 776-779 (1994) and Kent et al, U.S. patent 6,184,344. Briefly, in the approach a first peptide fragment is provided with an N-terminal cysteine having an unoxidized sulphydryl side chain, and a second peptide fragment is provided with a C-terminal thioester. The unoxidized sulphydryl side chain of the N-terminal cysteine is then condensed with the C-terminal thioester to produce an intermediate peptide fragment which links the first and second peptide fragments with a β -aminothioester bond. The β -aminothioester bond of the intermediate peptide fragment then undergoes an intramolecular rearrangement to produce the peptide fragment product which links the first and second peptide fragments with an amide bond. Preferably, the N-terminal cysteine of internal fragments are protected from undesired cyclization and/or concatenation reactions by a cyclic thiazolidine protecting group as described below. Preferably, such cyclic thiazolidine protecting group is a thioprolinyl group.

Peptide fragments having a C-terminal thioester may be produced as described in the following references, which are incorporated by reference: Kent et al, U.S. patent 6,184,344; Tam et al, Proc. Natl. Acad. Sci., 92: 12485-12489 (1995); Blake, Int. J. Peptide Protein Res., 17: 273 (1981); Canne et al, Tetrahedron Letters, 36: 1217-1220 (1995); Hackeng et al, Proc. Natl. Acad. Sci., 94: 7845-7850 (1997); or Hackeng et al, Proc. Natl. Acad. Sci., 96: 10068-10073 (1999). Preferably, the method described by Hackeng et al (1999) is employed. Briefly, peptide fragments are synthesized on a solid phase support (described below) typically on a 0.25 mmol scale by using the in situ neutralization/HBTU activation procedure for Boc chemistry disclosed by Schnolzer et al, Int. J. Peptide Protein Res., 40: 180-193 (1992), which reference is incorporated herein by reference. (HBTU is 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and Boc is tert-butoxycarbonyl). Each synthetic cycle consists of N^{α} -Boc removal by a 1- to 2-minute treatment with neat TFA, a 1-minute DMF flow wash, a 10- to 20-minute coupling time with 1.0 mmol of preactivated Boc-amino acid in the presence of DIEA, and a second DMF flow wash. (TFA is trifluoroacetic acid, DMF is N,N-dimethylformamide, and DIEA is N,N-diisopropylethylamine). N^{α} -Boc-amino acids (1.1 mmol) are preactivated for 3 minutes with 1.0 mmol of HBTU (0.5 M in DMF) in the presence of excess DIEA (3 mmol). After each coupling step, yields are determined by measuring residual free amine with a conventional quantitative ninhydrin assay, e.g. as disclosed in Sarin et al, Anal. Biochem., 117: 147-157 (1981). After coupling of Gln residues, a DCM flow wash is used before and after deprotection by using TFA, to prevent possible high-temperature (TFA/DMF)-catalyzed pyrrolidone formation. After chain assembly is completed, the peptide fragments are deprotected and cleaved from the resin by treatment with anhydrous HF for 1 hour at 0°C with 4% *p*-cresol as a scavenger. The imidazole side-chain 2,4-dinitrophenyl (dnp) protecting groups remain on the His residues because the dnp-removal procedure is incompatible with C-terminal thioester groups. However, dnp is gradually removed by thiols during the ligation reaction. After cleavage, peptide fragments are precipitated with ice-cold diethyl ether, dissolved in aqueous acetonitrile, and lyophilized.

Thioester peptide fragments described above are preferably synthesized on a trityl-associated mercaptopropionic acid-leucine (TAMPAL) resin, made as disclosed by Hackeng et al (1999), or comparable protocol. Briefly, N^α-Boc-Leu (4 mmol) is activated with 3.6 mmol of HBTU in the presence of 6 mmol of DIEA and coupled for 16 minutes to 2 mmol of p-methylbenzhydrylamine (MBHA) resin, or the equivalent. Next, 3 mmol of S-trityl mercaptopropionic acid is activated with 2.7 mmol of HBTU in the presence of 6 mmol of DIEA and coupled for 16 minutes to Leu-MBHA resin. The resulting TAMPAL resin can be used as a starting resin for polypeptide-chain assembly after removal of the trityl protecting group with two 1-minute treatments with 3.5% triisopropylsilane and 2.5% H₂O in TFA. The thioester bond can be formed with any desired amino acid by using standard in situ-neutralization peptide coupling protocols for 1 hour, as disclosed in Schnolzer et al (cited above). Treatment of the final peptide fragment with anhydrous HF yields the C-terminal activated mercaptopropionic acid-leucine (MPAL) thioester peptide fragments.

Preferably, thiazolidine-protected thioester peptide fragment intermediates are used in native chemical ligation under conditions as described by Hackeng et al (1999), or like conditions. Briefly, 0.1 M phosphate buffer (pH 8.5) containing 6 M guanidine, 4% (vol/vol) benzylmercaptan, and 4% (vol/vol) thiophenol is added to dry peptides to be ligated, to give a final peptide concentration of 1-3 mM at about pH 7, lowered because of the addition of thiols and TFA from the lyophilized peptide. Preferably, the ligation reaction is performed in a heating block at 37°C and is periodically vortexed to equilibrate the thiol additives. The reaction may be monitored for degree of completion by MALDI-MS or HPLC and electrospray ionization MS.

After a native chemical ligation reaction is completed or stopped, the N-terminal thiazolidine ring of the product is opened by treatment with a cysteine deprotecting agent, such as O-methylhydroxylamine (0.5 M) at pH 3.5-4.5 for 2 hours at 37°C, after which a 10-fold excess of Tris-(2-carboxyethyl)-phosphine is added to the reaction mixture to completely reduce any oxidizing reaction constituents prior to purification of the product by conventional preparative HPLC. Preferably, fractions containing the ligation product are identified by electrospray MS, are pooled, and lyophilized.

After the synthesis is completed and the final product purified, the final polypeptide product may be refolded by conventional techniques, e.g. Creighton, Meth. Enzymol., 107: 305-329 (1984); White, Meth. Enzymol., 11: 481-484 (1967); Wetlauffer, Meth. Enzymol., 107: 301-304 (1984), and the like. Preferably, a final product is refolded by air oxidation by the following, or like: The reduced lyophilized product is dissolved (at about 0.1 mg/mL) in 1 M guanidine hydrochloride (or like chaotropic agent) with 100 mM Tris, 10 mM methionine, at pH 8.6. After gentle overnight stirring, the re-folded product is isolated by reverse phase HPLC with conventional protocols.

Recombinant Expression Vectors and Host Cells

- The polynucleotide sequences described herein can be used in recombinant DNA molecules that direct the expression of the corresponding polypeptides in appropriate host cells. Because of the degeneracy in the genetic code, other DNA sequences may encode the equivalent amino acid sequence, and may be used to clone and express the Pro412-related proteins. Codons preferred by a particular host cell may be selected and substituted into the naturally occurring nucleotide sequences, to increase the rate and/or efficiency of expression. The nucleic acid (e.g., cDNA or genomic DNA) encoding the desired Pro412-related protein may be inserted into a replicable vector for cloning (amplification of the DNA), or for expression. The polypeptide can be expressed recombinantly in any of a number of expression systems according to methods known in the art (Ausubel, et al., editors, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1990). Appropriate host cells include yeast, bacteria, archebacteria, fungi, and insect and animal cells, including mammalian cells, for example primary cells, including stem cells, including, but not limited to bone marrow stem cells. More specifically, these include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors, and yeast transformed with yeast expression vectors. Also included, are insect cells infected with a recombinant insect virus (such as baculovirus), and mammalian expression systems. The nucleic acid sequence to be expressed may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.
- The Pro412-related proteins of the present invention are produced by culturing a host cell transformed with an expression vector containing a nucleic acid encoding a Pro412-related protein, under the appropriate conditions to induce or cause expression of the protein. The conditions appropriate for Pro412-related protein expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation. For example, the use of constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

A host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the

protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing, which cleaves a "prepro" form of the protein, may also be important for correct insertion, folding and/or function. By way of example, host cells such as CHO, HeLa, BHK, MDCK, 293, W138, etc. have specific cellular machinery

5 and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein. Of particular interest are *Drosophila melanogaster* cells, *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus subtilis*, SF9 cells, C129 cells, 293 cells, Neurospora, BHK, CHO, COS, and HeLa cells, fibroblasts, Schwannoma cell lines, immortalized mammalian myeloid and lymphoid cell lines,

10 Jukat cells, human cells and other primary cells.

The nucleic acid encoding a Pro412-related protein must be "operably linked" by placing it into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation.

15 Generally, "operably linked" DNA sequences are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic

20 oligonucleotide adaptors or linkers are used in accordance with conventional practice. Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention. The expression vector may comprise additional elements, for example, the expression vector may have

25 two replication systems, thus allowing it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most

30 Gram-negative bacteria, the 2: plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Further, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably, two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting

35 the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art.

Preferably, the expression vector contains a selectable marker gene to allow the selection

of transformed host cells. Selection genes are well known in the art and will vary with the host cell used. Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for Bacilli.

Host cells transformed with a nucleotide sequence encoding a Pro412-related protein may be cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein produced by a recombinant cell may be secreted, membrane-bound, or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides encoding the Pro412-related proteins can be designed with signal sequences which direct secretion of the Pro412-related proteins through a prokaryotic or eukaryotic cell membrane. The desired Pro412-related protein may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the Pro412-related proteins-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* a-factor leaders, the latter described in U.S. Pat. No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published Apr. 4, 1990), or the signal described in WO 90113646 published Nov. 15, 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders. According to the expression system selected, the coding sequence is inserted into an appropriate vector, which in turn may require the presence of certain characteristic "control elements" or "regulatory sequences." Appropriate constructs are known generally in the art (Ausubel, et al., 1990) and, in many cases, are available from commercial suppliers such as Invitrogen (San Diego, Calif.), Stratagene (La Jolla, Calif.), Gibco BRL (Rockville, Md.) or Clontech (Palo Alto, Calif.).

Expression in Bacterial Systems.

Transformation of bacterial cells may be achieved using an inducible promoter such as the hybrid lacZ promoter of the "BLUESCRIPT" Phagemid (Stratagene) or "pSPORT1" (Gibco BRL). In addition, a number of expression vectors may be selected for use in bacterial cells to produce cleavable fusion proteins that can be easily detected and/or purified, including, but not

limited to "BLUESCRIPT" (α -galactosidase; Stratagene) or pGEX (glutathione S-transferase; Promega, Madison, Wis.). A suitable bacterial promoter is any nucleic acid sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of the coding sequence of the Pro412-related gene into mRNA. A bacterial promoter has a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region typically includes an RNA polymerase binding site and a transcription initiation site. Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose and maltose, and sequences derived from biosynthetic enzymes such as tryptophan. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; for example, the tat promoter is a hybrid of the trp and lac promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. An efficient ribosome binding site is also desirable. The expression vector may also include a signal peptide sequence that provides for secretion of the Pro412-related protein in bacteria. The signal sequence typically encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell, as is well known in the art. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include drug resistance genes such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways. When large quantities of Pro412-related proteins are needed, e.g., for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the Pro412-related protein coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; PIN vectors [Van Heeke & Schuster *J Biol Chem* 264:5503-5509 (1989)]; PET vectors (Novagen, Madison Wis.); and the like. Expression vectors for bacteria include the various components set forth above, and are well known in the art. Examples include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others. Bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride mediated transfection, electroporation, and others.

Expression in Yeast.

Yeast expression systems are well known in the art, and include expression vectors for *Saccharomyces cerevisiae*, *Candida albicans* and *C. maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis* and *K. lactis*, *Pichia guilliermondii* and *P pastoris*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*. Examples of suitable promoters for use in yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., *J. Biol. Chem.* 255:2073 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.* 7:149 (1968); Holland, *Biochemistry* 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, alpha factor, the ADH2/GAPDH promoter, glucokinase alcohol oxidase, and PGH. [See, for example, Ausubel, et al., 1990; Grant et al., *Methods in Enzymology* 153:516-544, (1987)]. Other yeast promoters, which are inducible have the additional advantage of transcription controlled by growth conditions, include the promoter regions for alcohol dehydrogenase 2, isocytchrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657. Yeast selectable markers include ADE2, HIS4, LEU2, TRP1, and ALG7, which confers resistance to tunicamycin; the neomycin phosphotransferase gene, which confers resistance to G418; and the CUP1 gene, which allows yeast to grow in the presence of copper ions. Yeast expression vectors can be constructed for intracellular production or secretion of a Pro412-related protein from the DNA encoding the Pro412-related protein of interest. For example, a selected signal peptide and the appropriate constitutive or inducible promoter may be inserted into suitable restriction sites in the selected plasmid for direct intracellular expression of the Pro412-related protein. For secretion of the Pro412-related protein, DNA encoding the Pro412-related protein can be cloned into the selected plasmid, together with DNA encoding the promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (as needed), for expression of the Pro412-related protein. Yeast cells, can then be transformed with the expression plasmids described above, and cultured in an appropriate fermentation media. The protein produced by such transformed yeast can then be concentrated by precipitation with 10% trichloroacetic acid and analyzed following separation by SDS-PAGE and staining of the gels with Coomassie Blue stain. The recombinant Pro412-related proteins can subsequently be isolated and purified from the fermentation medium by techniques known to those of skill in the art.

35 *Expression in Mammalian Systems.*

The Pro412-related proteins may be expressed in mammalian cells. Mammalian expression systems are known in the art, and include retroviral vector mediated expression systems.

Mammalian host cells may be transformed with any of a number of different viral-based expression systems, such as adenovirus, where the coding region can be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome results in a viable virus capable of

5 expression of the polypeptide of interest in infected host cells. A preferred expression vector system is a retroviral vector system such as is generally described in PCT/US97/01019 and PCT/US97/101048. Suitable mammalian expression vectors contain a mammalian promoter which is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence for Pro412-related protein into mRNA. A

10 promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, using a located 25-30 base pairs upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element (enhancer element), typically located within 100 to 200 base pairs upstream of the TATA.

15 box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211, 504 published Jul. 5, 1989), adenovirus (such as Adenovirus 2), bovine

20 papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems. Transcription of a DNA encoding a Pro412-related protein by higher eukaryotes may be increased by inserting an enhancer sequence into the vector.

25 Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and

30 adenovirus enhancers. The enhancer is preferably located at a site 5' from the promoter. In general, the transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-translational cleavage and polyadenylation. Examples of transcription terminator and

35 polyadenylation signals include those derived from SV40. Long term, high-yield production of recombinant proteins can be effected in a stable expression system. Expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker

gene may be used for this purpose. Appropriate vectors containing selectable markers for use in mammalian cells are readily available commercially and are known to persons skilled in the art. Examples of such selectable markers include, but are not limited to herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase for use in tk- or hprt-cells, respectively. The 5 methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, viral infection, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

10

Expression in Insect Cells.

Pro412-related proteins may also be produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art. In one such system, the Pro412-related proteins-encoding DNA is fused upstream 15 of an epitope tag contained within a baculovirus expression vector. *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* Sf9 cells or in *Trichoplusia* larvae. The Pro412-related proteins-encoding sequence is cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of a Pro412-related proteins-encoding sequence will 20 render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which the Pro412-related proteins is expressed [Smith et al., J. Wol. 46:584 (1994); Engelhard E K et al., Proc. Nat. Acad. Sci. 91:3224-3227 (1994)]. Suitable epitope tags for fusion to the Pro412-related proteins-encoding DNA include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A 25 variety of plasmids may be employed, including commercially available plasmids such as pVL1393 (Novagen). Briefly, the Pro412-related proteins-encoding DNA or the desired portion of the Pro412-related proteins-encoding DNA is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking restriction sites. The PCR product is then digested with the selected restriction enzymes and subcloned into an expression vector.

30 Recombinant baculovirus is generated by co-transfected the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL), or other methods known to those of skill in the art. Virus is produced by day 4-5 of culture in Sf9 cells at 28°C., and used for further amplifications. Procedures are performed as further described in O'Reilly et al., *BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL*, Oxford University Press (1994). Extracts 35 may be prepared from recombinant virus-infected Sf9 cells as described in Rupert et al., *Nature* 362:175-179 (1993). Alternatively, expressed epitope-tagged Pro412-related proteins can be

purified by affinity chromatography, or for example, purification of an IgG tagged (or Fc tagged) Pro412-related protein can be performed using chromatography techniques, including Protein A or protein G column chromatography.

5 Evaluation of Gene Expression.

- Gene expression may be evaluated in a sample directly, for example, by standard techniques known to those of skill in the art, e.g., Southern blotting for DNA detection, Northern blotting to determine the transcription of mRNA, dot blotting (DNA or RNA), or in situ hybridization, using an appropriately labeled probe, based on the sequences provided herein.
- 10 Alternatively, antibodies may be used in assays for detection of nucleic acids, such as specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. Such antibodies may be labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. Gene expression, alternatively, may be measured by
- 15 immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to directly evaluate the expression of Pro412-related proteins. Antibodies useful for such immunological assays may be either monoclonal or polyclonal, and may be prepared against a native sequence Pro412-related proteins based on the DNA sequences provided herein.

20 Purification of Expressed Protein.

- Expressed Pro412-related proteins may be purified or isolated after expression, using any of a variety of methods known to those skilled in the art. The appropriate technique will vary depending upon what other components are present in the sample. Contaminant components that are removed by isolation or purification are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other solutes. The purification step(s) selected will depend, for example, on the nature of the production process used and the particular Pro412-related protein produced. A Pro412-related protein or protein may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Alternatively, cells employed in expression of Pro412-related proteins can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or by use of cell lysing agents. Exemplary purification methods include, but are not limited to, ion-exchange column chromatography; chromatography using silica gel or a cation-exchange resin such as DEAE; gel filtration using, for example, Sephadex G-75, protein A Sepharose columns to remove contaminants such as IgG; chromatography using metal chelating columns to bind epitope-tagged forms of the Pro412-related protein; ethanol precipitation; reverse phase HPLC, chromatofocusing; SDS-PAGE; and

ammonium sulfate precipitation. Ordinarily, an isolated Pro412-related protein will be prepared by at least one purification step. For example, the Pro412-related protein may be purified using a standard anti-Pro412-related antibody column. Ultrafiltration and dialysis techniques, in conjunction with protein concentration, are also useful (see, for example, Scopes, R., *PROTEIN*

- 5 *PURIFICATION*, Springer-Verlag, New York, N.Y., 1982). The degree of purification necessary will vary depending on the use of the Pro412-related proteins. In some instances no purification will be necessary. Once expressed and purified as needed, the Pro412-related proteins and nucleic acids of the present invention are useful in a number of applications, as detailed below.

10 *Labeling of Expressed Protein.*

- The nucleic acids, proteins and antibodies of the invention may be labeled. By labeled herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) immune labels, which may be antibodies or 15 antigens; and c) colored or fluorescent dyes. The labels may be incorporated into the compound at any position that does not interfere with the biological activity or characteristic of the compound which is being detected.

Pro412-related fusion Proteins.

- 20 The Pro412-related proteins of the present invention may also be modified in a way to form chimeric molecules comprising a Pro412-related protein fused to another, heterologous polypeptide or amino acid sequence. The term "fusion protein" used herein includes but is not limited to a chimeric polypeptide comprising a Pro412-related protein, or domain sequence thereto, fused to a "targeting polypeptide". The targeting polypeptide has enough residues to facilitate targeting to a particular cell type or receptor, yet is short enough such that it does not interfere with the biological function of the Pro412-related protein. The targeting polypeptide preferably is also fairly unique so that the fusion protein does not substantially cross-react with other cell types or receptors. Suitable targeting polypeptides generally have at least about 10 amino acid residues and usually between from about 10 to about 500 amino acid residues. Preferred targeting polypeptides 25 have from about 20 to about 200 amino acid residues. The fusion protein may also comprise a fusion of a Pro412-related protein with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the Pro412-related proteins. Such epitope-tagged forms of a Pro412-related protein can be detected using an antibody against the tag polypeptide. Also, provision of the 30 epitope tag enables a Pro412-related protein to be readily purified by using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Alternatively, the fusion protein may comprise a fusion of a Pro412-related protein with an immunoglobulin or a particular region of an

immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule or, for example, GM-CSF. Preferred fusion proteins include, but are not limited to, molecules that facilitate immune targeting of a Pro412-related protein. The Pro412-related fusion protein may be made for various other purposes using techniques well known in the art. For example, for the creation of antibodies, if the desired epitope is small, a partial or complete Pro412-related protein may be fused to a carrier protein to form an immunogen. Alternatively, the Pro412-related protein may be made as a fusion protein to increase the ability of the antigen to stimulate cellular and/or humoral (antibody-based) immune responses, or for other reasons.

10 *Synthetic Genes for Pro412-related proteins.*

Once nucleic acid sequence and/or amino acid sequence information is available for a native protein a variety of techniques become available for producing virtually any mutation in the native sequence, e.g. Shortle, in *Science*, Vol. 229, pgs. 1193-1201 (1985); Zoller and Smith, *Methods in Enzymology*, Vol. 100, pgs. 468-500 (1983); Mark et al., U.S. Patent 4,518,584; Wells et al., in *Gene*, Vol. 34, pgs. 315-323 (1985); Estell et al., *Science*, Vol. 233, pgs. 659-663 (1986); Mullenbach et al., *J. Biol. Chem.*, Vol. 261, pgs. 719-722 (1986), and Feretti et al., *Proc. Natl. Acad. Sci.*, Vol. 83, pgs. 597-603 (1986). Accordingly, these references are incorporated by reference.

20 Variants of the natural polypeptide (sometime referred to as "muteins") may be desirable in a variety of circumstances. For example, undesirable side effects might be reduced by certain variants, particularly if the side effect activity is associated with a different part of the polypeptide from that of the desired activity. In some expression systems, the native polypeptide may be susceptible to degradation by proteases. In such cases, selected substitutions and/or deletions of amino acids which change the susceptible sequences can significantly enhance yields, e.g. British patent application 2173-804-A where Arg at position 275 of human tissue plasminogen activator is replaced by Gly or Glu. Variants may also increase yields in purification procedures and/or increase shelf lives of proteins by eliminating amino acids susceptible to oxidation, acylation, alkylation, or other chemical modifications. For example, methionines readily undergo oxidation to form sulfoxides, which in many proteins is associated with loss of biological activity, e.g. Brot and Weissbach, *Arch. Biochem. Biophys.*, Vol. 223, pg. 271 (1983). Often methionines can be replaced by more inert amino acids with little or no loss of biological activity, e.g. Australian patent application AU-A-52451/86. In bacterial expression systems, yields can sometimes be increased by eliminating or replacing conformationally inessential cysteine residues, e.g. Mark et al., U.S. Patent 4,518,584.

35 Preferably cassette mutagenesis is employed to generate mutant proteins. A synthetic gene is constructed with a sequence of unique (when inserted in an appropriate vector) restriction endonuclease sites spaced approximately uniformly along the gene. The unique restriction sites

allow segments of the gene to be conveniently excised and replaced with synthetic oligonucleotides (i.e. "cassettes") which code for desired mutations. Determination of the number and distribution of unique restriction sites entails the consideration of several factors including (1) preexisting restriction sites in the vector to be employed in expression, (2) whether species or 5 genera-specific codon usage is desired, (3) the number of different non-vector-cutting restriction endonucleases available (and their multiplicities within the synthetic gene), and (4) the convenience and reliability of synthesizing and/or sequencing the segments between the unique restriction sites.

The above technique is a convenient way to effect conservative amino acid substitutions, and the 10 like, in the native protein sequence. "Conservative" as used herein means (i) that the alterations are as conformationally neutral as possible, that is, designed to produce minimal changes in the tertiary structure of the mutant polypeptides as compared to the native protein, and (ii) that the alterations are as antigenically neutral as possible, that is, designed to produce minimal changes in the antigenic determinants of the mutant polypeptides as compared to the native protein. . The 15 following is a preferred categorization of amino acids into similarity classes: aromatic (phe, trp, tyr), hydrophobic (leu, ile, val), polar (gln, asn), basic (arg, lys, his), acidic (asp, glu), small (ala, ser, thr, met, gly). Conformational neutrality is desirable for preserving biological activity, and antigenic neutrality is desirable for avoiding the triggering of immunogenic responses in patients or animals treated with the compounds of the invention. While it is difficult to select with absolute 20 certainty which alternatives will be conformationally and antigenically neutral, rules exist which can guide those skilled in the art to make alterations that have high probabilities of being conformationally and antigenically neutral, e.g. Anfisen (cited above); Berzofsky, Science, Vol. 229, pgs. 932-940 (1985); and Bowie et al, Science, Vol. 247, pgs. 1306-1310 (1990). Some of the more important rules include (1) substitution of hydrophobic residues are less likely to produce 25 changes in antigenicity because they are likely to be located in the protein's interior, e.g. Berzofsky (cited above) and Bowie et al (cited above); (2) substitution of physiochemically similar, i.e. synonymous, residues are less likely to produce conformational changes because the replacement amino acid can play the same structural role as the substituted amino acid; and (3) alteration of evolutionarily conserved sequences is likely to produce deleterious conformational effects because 30 evolutionary conservation suggests sequences may be functionally important. In addition to such basic rules for selecting variant sequences, assays are available to confirm the biological activity and conformation of the engineered molecules. Biological assays for the polypeptides of the invention are described more fully above. Changes in conformation can be tested by at least two well known assays: the microcomplement fixation method, e.g. Wasserman et al., J. Immunol., 35 Vol. 87, pgs. 290-295 (1961), or Levine et al. Methods in Enzymology, Vol. 11, pgs. 928-936 (1967) used widely in evolutionary studies of the tertiary structures of proteins; and affinities to sets of conformation-specific monoclonal antibodies, e.g. Lewis et al., Biochemistry, Vol. 22, pgs.

948-954 (1983).

Transgenic animals

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which Pro412-related proteins-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous Pro412-related proteins sequences have been introduced into their genome or homologous recombinant animals in which endogenous Pro412-related protein sequences have been altered. Such animals are useful for studying the function and/or activity of a Pro412-related protein or fragment thereof and for identifying and/or evaluating modulators of Pro412-related activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous Pro412-related gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a Pro412-related proteins-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection or retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The Pro412-related proteins cDNA sequence or a fragment thereof can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human Pro412-related gene, such as a mouse or rat Pro412-related gene, can be used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a Pro412-related proteins transgene to direct expression of a Pro412-related protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Pat. No. 4,873,191 by Wagner et al. and in Hogan, B., Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986, the disclosure of which is incorporated herein by reference in its entirety). Similar methods are used for production of other

transgenic animals. A transgenic founder animal can be identified based upon the presence of a Pro412-related protein transgene in its genome and/or expression of Pro412-related proteins mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene 5 encoding a Pro412-related protein can further be bred to other transgenic animals carrying other transgenes.

To create an animal in which a desired nucleic acid has been introduced into the genome via homologous recombination, a vector is prepared which contains at least a portion of a Pro412-related gene into which a deletion, addition or substitution has been introduced to thereby alter, 10 e.g., functionally disrupt, the Pro412-related gene. The Pro412-related gene can be a human gene, but more preferably, is a non-human homologue of a human Pro412-related gene (e.g., a cDNA isolated by stringent hybridization with a nucleotide sequence coding for a Pro412-related protein). For example, a mouse Pro412-related gene can be used to construct a homologous recombination vector suitable for altering an endogenous gene in the mouse genome. In a preferred embodiment, 15 the vector is designed such that, upon homologous recombination, the endogenous Pro412-related gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous Pro412-related gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the 20 expression of the endogenous Pro412-related protein). In the homologous recombination vector, the altered portion of the Pro412-related protein or gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the Pro412-related gene to allow for homologous recombination to occur between the exogenous Pro412-related gene carried by the vector and an endogenous Pro412-related gene in an embryonic stem cell. The additional flanking Pro412-related 25 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas, K. R. and Capecchi, M. R. (1987) Cell 51:503, the disclosure of which is incorporated herein by reference in its entirety, for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line 30 (e.g., by electroporation) and cells in which the introduced Pro412-related gene has homologously recombined with the endogenous Pro412-related gene are selected (see e.g., Li, E. et al. (1992) Cell 69:915, the disclosure of which is incorporated herein by reference in its entirety). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley, A. in Teratocarcinomas and Embryonic Stem Cells. A Practical 35 Approach, E. J. Robertson, ed (IRL, Oxford, 1987) pp. 113-152, the disclosure of which is incorporated herein by reference in its entirety). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring

the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, A. (1991) Current Opinion in Biotechnology 5 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec et al.; WO 91/01140 by Smithies et al.; WO 92/0968 by Zijlstra et al.; and WO 93/04169 by Berns et al., the disclosures of which are incorporated herein by reference in their entireties.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso et al. (1992) PNAS 89:6232-6236, the disclosure of which is incorporated herein by reference in its entirety. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) Science 251:1351-1355, the disclosure of which is incorporated herein by reference in its entirety). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

20

Anti-Pro412-related antibodies

The present invention provides antibodies and binding compositions specific for Pro412-related proteins. Such antibodies and binding compositions include polyclonal antibodies, monoclonal antibodies, Fab and single chain Fv fragments thereof, bispecific antibodies, heteroconjugates, humanized antibodies, and the like. Such antibodies and binding compositions may be produced in a variety of ways, including hybridoma cultures, recombinant expression in bacteria or mammalian cell cultures, recombinant expression in transgenic animals, and the like. There is abundant guidance in the literature for selecting a particular production methodology, e.g. Chadd and Chamow, *Curr. Opin. Biotechnol.*, 12: 188-194 (2001).

30

The choice of manufacturing methodology depends on several factors including the antibody structure desired, the importance of carbohydrate moieties on the antibodies, ease of culturing and purification, cost, and the like. Many different antibody structures may be generated using standard expression technology, including full-length antibodies, antibody fragments, such as Fab and Fv fragments, as well as chimeric antibodies comprising components from different species. Antibody fragments of small size, such as Fab and Fv fragments, having no effector functions and limited pharmacokinetic activity may be generated in a bacterial expression system. Single chain Fv fragments are highly selective for *in vivo* tumors, show good tumor penetration

and low immunogenicity, and are cleared rapidly from the blood, e.g. Freyre et al, *J. Biotechnol.*, 76: 157-163 (2000). Thus, such molecules are desirable for radioimmuno detection and in situ radiotherapy. Whenever pharmacokinetic activity in the form of increased half-life is required for therapeutic purposes, then a full-length antibodies are preferable. For example, immunoglobulin G (IgG) the molecule may be one of four subclasses: $\gamma 1$, $\gamma 2$, $\gamma 3$, or $\gamma 4$. If a full-length antibody with effector function is required, then IgG subclasses $\gamma 1$ or $\gamma 3$ are preferred, and IgG subclass $\gamma 1$ is most preferred. The $\gamma 1$ and $\gamma 3$ subclasses exhibit potent effector function, complement activation, and promote antibody-dependent cell-mediated cytotoxicity through interaction with specific Fc receptors, e.g. Raju et al, *Glycobiology*, 10: 477-486 (2000); Lund et al, *J. Immunol.*, 147: 2657-10 2662 (1991).

Polyclonal Antibodies.

The anti-Pro412-related antibodies of the present invention may be polyclonal antibodies. Such polyclonal antibodies can be produced in a mammal, for example, following one or more 15 injections of an immunizing agent, and preferably, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected into the mammal by a series of subcutaneous or intraperitoneal injections. The immunizing agent may include Pro412-related proteins or a fusion protein thereof. It may be useful to conjugate the antigen to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include, but are not limited to, keyhole 20 limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Adjuvants include, for example, Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicoryno-mycolate). The immunization protocol may be determined by one skilled in the art based on standard protocols or by routine experimentation.

25

Monoclonal Antibodies.

Alternatively, the anti-Pro412-related antibodies may be monoclonal antibodies. Monoclonal antibodies may be produced by hybridomas, wherein a mouse, hamster, or other appropriate host animal, is immunized with an immunizing agent to elicit lymphocytes that 30 produce or are capable of producing antibodies that will specifically bind to the immunizing agent, e.g. Kohler and Milstein, *Nature* 256:495 (1975). Alternatively, the lymphocytes may be immunized in vitro. The immunizing agent will typically include the Pro412-related proteins or a fusion protein thereof. Generally, spleen cells or lymph node cells are used if non-human mammalian sources are desired, or peripheral blood lymphocytes ("PBLs") are used if cells of 35 human origin. The lymphocytes are fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to produce a hybridoma cell, e.g. Goding, *MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE*, Academic Press, pp 59-103 (1986); Liddell and

Cryer, A Practical Guide to Monoclonal Antibodies (John Wiley & Sons, New York, 1991); Malik and Lillenq, Editors, Antibody Techniques (Academic Press, New York, 1994). In general, immortalized cell lines are transformed mammalian cells, for example, myeloma cells of rat, mouse, bovine or human origin. The hybridoma cells are cultured in a suitable culture medium that

5 preferably contains one or more substances that inhibit the growth or survival of unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine (HAT), substances which prevent the growth of HGPRT-deficient cells. Preferred immortalized cell lines are those that fuse efficiently, support

10 stable high level production of antibody, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine or human myeloma lines, which can be obtained, for example, from the American Type Culture Collection (ATCC), Rockville, MD. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies, e.g. Kozbor, *J. Immunol.* 133:3001 (1984); Brodeur

15 et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, pp. 51-63 (1987).

The culture medium (supernatant) in which the hybridoma cells are cultured can be assayed for the presence of monoclonal antibodies directed against a Pro412-related protein. Preferably, the binding specificity of monoclonal antibodies present in the hybridoma supernatant

20 is determined by immunoprecipitation or by an in vitro binding assay, such as radio-immunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Appropriate techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, *Anal. Biochem.* 107:220 (1980). After the desired antibody-producing hybridoma cells are identified, the cells may be cloned by limiting

25 dilution procedures and grown by standard methods [Goding, 1986]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown in vivo as ascites in a mammal. The monoclonal antibodies secreted by selected clones may be isolated or purified from the culture medium or

30 ascites fluid by immunoglobulin purification procedures routinely used by those of skill in the art such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Pat. No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be isolated from the Pro412-related proteins-specific hybridoma cells and sequenced,

35 e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies. Once isolated, the DNA may be inserted into an expression vector, which is then transfected into host cells such as simian COS cells, Chinese

- hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for the human heavy and light chain constant domains for the homologous murine sequences [Morrison et al., *Proc. Nat.*
- 5 *Acad. Sci.* 81:6851-6855 (1984); Neuberger et al., *Nature* 312:604-608 (1984); Takeda et al., *Nature* 314:452-454 (1985)], or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. The non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention
- 10 to create a chimeric bivalent antibody. The antibodies may also be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, in vitro methods are suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.
- 15 Antibodies and antibody fragments characteristic of hybridomas of the invention can also be produced by recombinant means by extracting messenger RNA, constructing a cDNA library, and selecting clones which encode segments of the antibody molecule. The following are exemplary references disclosing recombinant techniques for producing antibodies: Wall et al., *Nucleic Acids Research*, Vol. 5, pgs. 3113-3128 (1978); Zukut et al., *Nucleic Acids Research*, Vol.
- 20 8, pgs. 3591-3601 (1980); Cabilly et al., *Proc. Natl. Acad. Sci.*, Vol. 81, pgs. 3273-3277 (1984); Boss et al., *Nucleic Acids Research*, Vol. 12, pgs. 3791-3806 (1984); Amster et al., *Nucleic Acids Research*, Vol. 8, pgs. 2055-2065 (1980); Moore et al., U.S. Patent 4,642,334; Skerra et al., *Science*, Vol. 240, pgs. 1038-1041 (1988); Huse et al., *Science*, Vol. 246, pgs. 1275-1281 (1989); and U.S. patents 6,054,297; 5,530,101; 4,816,567; 5,750,105; and 5,648,237; which patents are
- 25 incorporated by reference. In particular, such techniques can be used to produce interspecific monoclonal antibodies, wherein the binding region of one species is combined with non-binding region of the antibody of another species to reduce immunogenicity, e.g. Liu et al., *Proc. Natl. Acad. Sci.*, Vol. 84, pgs. 3439-3443 (1987), and patents 6,054,297 and 5,530,101. Preferably, recombinantly produced Fab and Fv fragments are expressed in bacterial host systems. Preferably,
- 30 full-length antibodies are produced by mammalian cell culture techniques. More preferably, full-length antibodies are expressed in Chinese Hamster Ovary (CHO) cells or NSO cells.
- Both polyclonal and monoclonal antibodies can be screened by ELISA. As in other solid phase immunoassays, the test is based on the tendency of macromolecules to adsorb nonspecifically to plastic. The irreversibility of this reaction, without loss of immunological activity, allows the formation of antigen-antibody complexes with a simple separation of such complexes from unbound material. To titrate antipeptide serum, peptide conjugated to a carrier different from that used in immunization is adsorbed to the wells of a 96-well microtiter plate. The

adsorbed antigen is then allowed to react in the wells with dilutions of anti-peptide serum. Unbound antibody is washed away, and the remaining antigen-antibody complexes are allowed to react with antibody specific for the IgG of the immunized animal. This second antibody is conjugated to an enzyme such as alkaline phosphatase. A visible colored reaction product produced when the enzyme substrate is added indicates which wells have bound anti-peptide antibodies. The use of spectrophotometer readings allows better quantification of the amount of peptide-specific antibody bound. High-titer antisera yield a linear titration curve between 10^{-3} and 10^{-5} dilutions.

10 *Pro412-related antibodies.*

The invention includes peptides derived from Pro412-related proteins, and immunogens comprising conjugates between carriers and peptides of the invention. The term immunogen as used herein refers to a substance which is capable of causing an immune response. The term carrier as used herein refers to any substance which when chemically conjugated to a peptide of the invention permits a host organism immunized with the resulting conjugate to generate antibodies specific for the conjugated peptide. Carriers include red blood cells, bacteriophages, proteins, or synthetic particles such as agarose beads. Preferably, carriers are proteins, such as serum albumin, gamma-globulin, keyhole limpet hemocyanin, thyroglobulin, ovalbumin, fibrinogen, or the like.

20 The general technique of linking synthetic peptides to a carrier is described in several references, e.g. Walter and Doolittle, "Antibodies Against Synthetic Peptides," in Setlow et al., eds., *Genetic Engineering*, Vol. 5, pgs. 61-91 (Plenum Press, N.Y., 1983); Green et al. *Cell*, Vol. 28, pgs. 477-487 (1982); Lerner et al., *Proc. Natl. Acad. Sci.*, Vol. 78, pgs. 3403-3407 (1981); Shimizu et al., U.S. Patent 4,474,754; and Ganfield et al., U.S. Patent 4,311,639. Accordingly, 25 these references are incorporated by reference. Also, techniques employed to link haptens to carriers are essentially the same as the above-referenced techniques, e.g. chapter 20 in Tijssen *Practice and Theory of Enzyme Immunoassays* (Elsevier, New York, 1985). The four most commonly used schemes for attaching a peptide to a carrier are (1) glutaraldehyde for amino coupling, e.g. as disclosed by Kagan and Glick, in Jaffe and Behrman, eds. *Methods of Hormone* 30 *Radioimmunoassay*, pgs. 328-329 (Academic Press, N.Y., 1979), and Walter et al. *Proc. Natl. Acad. Sci.*, Vol. 77, pgs. 5197-5200 (1980); (2) water-soluble carbodiimides for carboxyl to amino coupling, e.g. as disclosed by Hoare et al., *J. Biol. Chem.*, Vol. 242, pgs. 2447-2453 (1967); (3) bis-diazobenzidine (DBD) for tyrosine to tyrosine sidechain coupling, e.g. as disclosed by Bassiri et al., pgs. 46-47, in Jaffe and Behrman, eds. (cited above), and Walter et al. (cited above); and (4) 35 maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) for coupling cysteine (or other sulphydryls) to amino groups, e.g. as disclosed by Kitagawa et al., *J. Biochem. (Tokyo)*, Vol. 79, pgs. 233-239 (1976), and Lerner et al. (cited above). A general rule for selecting an appropriate method for

coupling a given peptide to a protein carrier can be stated as follows: the group involved in attachment should occur only once in the sequence, preferably at the appropriate end of the segment. For example, BDB should not be used if a tyrosine residue occurs in the main part of a sequence chosen for its potentially antigenic character. Similarly, centrally located lysines rule out
5 the glutaraldehyde method, and the occurrences of aspartic and glutamic acids frequently exclude the carbodiimide approach. On the other hand, suitable residues can be positioned at either end of chosen sequence segment as attachment sites, whether or not they occur in the "native" protein sequence. Internal segments, unlike the amino and carboxy termini, will differ significantly at the "unattached end" from the same sequence as it is found in the native protein where the polypeptide
10 backbone is continuous. The problem can be remedied, to a degree, by acetylating the α -amino group and then attaching the peptide by way of its carboxy terminus. The coupling efficiency to the carrier protein is conveniently measured by using a radioactively labeled peptide, prepared either by using a radioactive amino acid for one step of the synthesis or by labeling the completed peptide by the iodination of a tyrosine residue. The presence of tyrosine in the peptide also allows
15 one to set up a sensitive radioimmune assay, if desirable. Therefore, tyrosine can be introduced as a terminal residue if it is not part of the peptide sequence defined by the native polypeptide.

Preferred carriers are proteins, and preferred protein carriers include bovine serum albumin, myoglobin, ovalbumin (OVA), keyhole limpet hemocyanin (KLH), or the like. Peptides can be linked to KLH through cysteines by MBS as disclosed by Liu et al., Biochemistry,
20 Vol. 18, pgs. 690-697 (1979). The peptides are dissolved in phosphate-buffered saline (pH 7.5), 0.1 M sodium borate buffer (pH 9.0) or 1.0 M sodium acetate buffer (pH 4.0). The pH for the dissolution of the peptide is chosen to optimize peptide solubility. The content of free cysteine for soluble peptides is determined by Ellman's method, Ellman, Arch. Biochem. Biophys., Vol. 82, pg. 7077 (1959). For each peptide, 4 mg KLH in 0.25 ml of 10 mM sodium phosphate buffer (pH 7.2)
25 is reacted with 0.7 mg MBS (dissolved in dimethyl formamide) and stirred for 30 min at room temperature. The MBS is added dropwise to ensure that the local concentration of formamide is not too high, as KLH is insoluble in >30% formamide. The reaction product, KLH-MBS, is then passed through Sephadex G-25 equilibrated with 50 mM sodium phosphate buffer (pH 6.0) to remove free MBS, KLH recovery from peak fractions of the column eluate (monitored by OD280)
30 is estimated to be approximately 80%. KLH-MBS is then reacted with 5 mg peptide dissolved in 1 ml of the chosen buffer. The pH is adjusted to 7.75 and the reaction is stirred for 3 hr at room temperature. Coupling efficiency is monitored with radioactive peptide by dialysis of a sample of the conjugate against phosphate-buffered saline, and ranged from 8% to 60%. Once the peptide-carrier conjugate is available polyclonal or monoclonal antibodies are produced by
35 standard techniques, e.g. as disclosed by Campbell, Monoclonal Antibody Technology (Elsevier, New York, 1984); Hurrell, ed. Monoclonal Hybridoma Antibodies: Techniques and Applications (CRC Press, Boca Raton, FL, 1982); Schreier et al. Hybridoma Techniques (Cold Spring Harbor

Laboratory, New York, 1980); U.S. Patent 4,562,003; or the like. In particular, U.S. Patent 4,562,003 is incorporated by reference.

Humanized Antibodies.

- 5 The anti-Pro412-related antibodies of the invention may further comprise humanized antibodies or human antibodies. The term "humanized antibody" refers to humanized forms of non-human (e.g., murine) antibodies that are chimeric antibodies, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab'), or other antigen-binding partial sequences of antibodies) which contain some portion of the sequence derived from non-human antibody.
- 10 Humanized antibodies include human immunoglobulins in which residues from a complementary determining region (CDR) of the human immunoglobulin are replaced by residues from a CDR of a non-human species such as mouse, rat or rabbit having the desired binding specificity, affinity and capacity. In general, the humanized antibody will comprise substantially all of at least one, and generally two, variable domains, in which all or substantially all of the CDR regions correspond to
- 15 those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature* 321:522-525 (1986) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)]. Methods for humanizing non-human antibodies are well known in the art.
- 20 Generally, a humanized antibody has one or more amino acids introduced into it from a source which is non-human in order to more closely resemble a human antibody, while still retaining the original binding activity of the antibody. Methods for humanization of antibodies are further detailed in Jones et al., *Nature* 321:522-525 (1986); Ricchmann et al., *Nature* 332:323-327 (1988); and Verhoeyen et al., *Science* 239:1534-1536 (1988). Such "humanized" antibodies are chimeric
- 25 antibodies in that substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species.

Heteroconjugate Antibodies.

- 30 Heteroconjugate antibodies which comprise two covalently joined antibodies, are also within the scope of the present invention. Heteroconjugate antibodies may be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be prepared using a disulfide exchange reaction or by forming a thioether bond.

Bispecific Antibodies.

- 35 Bispecific antibodies have binding specificities for at least two different antigens. Such antibodies are monoclonal, and preferably human or humanized. One of the binding specificities of

a bispecific antibody of the present invention is for a Pro412-related proteins, and the other one is preferably for a cell-surface protein or receptor or receptor subunit. Methods for making bispecific antibodies are known in the art, and in general, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs in hybridoma cells, where the two heavy chains have different specificities, e.g. Milstein and Cuello, *Nature* 305:537-539 (1983). Given that the random assortment of immunoglobulin heavy and light chains results in production of potentially ten different antibody molecules by the hybridomas, purification of the correct molecule usually requires some sort of affinity purification, e.g. affinity chromatography.

10

Antibody antagonists.

Preferrably, antagonists of the invention are derived from antibodies specific for Pro412-related proteins. More preferably, the antagonists of the invention comprise fragments or binding compositions specific for Pro412-related proteins. Antibodies comprise an assembly of polypeptide chains linked together by disulfide bridges. Two major polypeptide chains, referred to as the light chain and the heavy chain, make up all major structural classes (isotypes) of antibody. Both heavy chains and light chains are further divided into subregions referred to as variable regions and constant regions. Heavy chains comprise a single variable region and three different constant regions, and light chains comprise a single variable region (different from that of the heavy chain) and a single constant region (different from those of the heavy chain). The variable regions of the heavy chain and light chain are responsible for the antibody's binding specificity. As used herein, the term "heavy chain variable region" means a polypeptide (1) which is from 110 to 125 amino acids in length, and (2) whose amino acid sequence corresponds to that of a heavy chain of a monoclonal antibody of the invention, starting from the heavy chain's N-terminal amino acid. Likewise, the term "light chain variable region" means a polypeptide (1) which is from 95 to 115 amino acids in length, and (2) whose amino acid sequence corresponds to that of a light chain of a monoclonal antibody of the invention, starting from the light chain's N-terminal amino acid. As used herein the term "monoclonal antibody" refers to homogeneous populations of immunoglobulins which are capable of specifically binding to Pro412-related proteins. As used herein the term "binding composition" means a composition comprising two polypeptide chains (1) which, when operationally associated, assume a conformation having high binding affinity for Pro412-related proteins, and (2) which are derived from a hybridoma producing monoclonal antibodies specific for Pro412-related proteins. The term "operationally associated" is meant to indicate that the two polypeptide chains can be positioned relative to one another for binding by a variety of means, including by association in a native antibody fragment, such as Fab or Fv, or by way of genetically engineered cysteine-containing peptide linkers at the carboxyl termini. Normally, the two polypeptide chains correspond to the light chain variable region and heavy chain

variable region of a monoclonal antibody specific for Pro412-related proteins. Preferably, antagonists of the invention are derived from monoclonal antibodies specific for Pro412-related proteins. Monoclonal antibodies capable of blocking, or neutralizing, Pro412-related proteins are selected by their ability to inhibit Pro412-related proteins-induced effects.

5 The use and generation of fragments of antibodies is also well known, e.g. Fab fragments: Tijssen, Practice and Theory of Enzyme Immunoassays (Elsevier, Amsterdam, 1985); and Fv fragments: Hochman et al. Biochemistry, Vol. 12, pgs. 1130-1135 (1973), Sharon et al., Biochemistry, Vol. 15, pgs. 1591-1594 (1976) and Ehrlich et al., U.S. Patent 4,355,023; and antibody half molecules: Auditore- Hargreaves, U.S. Patent 4,470,925.

10 Preferably, monoclonal antibodies, Fv fragments, Fab fragments, or other binding compositions derived from monoclonal antibodies of the invention have a high affinity to Pro412-related proteins. The affinity of monoclonal antibodies and related molecules to Pro412-related proteins may be measured by conventional techniques including plasmon resonance, ELISA, equilibrium dialysis, and the like. Affinity measurement by plasmon resonance techniques may be carried out, for example, using a BIACore 2000 instrument (Biacore AB, Uppsala, Sweden) in accordance with the manufacturer's recommended protocol. Preferably, is measured by ELISA, for example, as described in U.S. patent 6,235,883, or like reference. Preferably, the dissociation constant between Pro412-related proteins and monoclonal antibodies of the invention is less than 10^{-5} molar. More preferably, such dissociation constant is less than 10^{-8} molar; still more preferably, such dissociation constant is less than 10^{-9} molar; and most preferably, such dissociation constant is in the range of 10^{-9} to 10^{-11} molar.

20 The invention includes kits for performing either quantitative immunoassays or PCRs to detect the abundances of Pro412-related proteins or its mRNA transcripts, respectively. Preferably, monoclonal antibody kits of the invention includes at least one monoclonal antibody specific for Pro412-related proteins. Kits of the invention further include labeling means to detect the quantity of each monoclonal antibody bound to its target protein. Preferably, quantitative measurements are made using a conventional ELISA format, as described more fully below.

25 The antibodies of the present invention find use in diagnostic assays for the determination of Pro412-related proteins. The antibodies of the invention may be used in most assays involving antigen-antibody reactions. The assays may be homogeneous or heterogeneous. In a homogeneous assay approach, the sample can be a biological sample or fluid such as serum, urine, whole blood, lymphatic fluid, plasma, saliva, and the like, cells, tissue, and material secreted by cells or tissues cultured in vitro. The sample can be pretreated if necessary to remove unwanted materials. The immunological reaction usually involves the specific antibody, labeled analyte, and the sample suspected of containing the analyte. The analyte can be directly labeled with the label or indirectly labeled with a means for incorporating the label such as conjugation of the analyte to biotin and having labeled avidin or anti-biotin. The signal from the label is modified, directly or indirectly,

- upon the binding of the antibody of the labeled analyte. Both the immunological reaction and detection of the extent thereof are carried out in a homogeneous solution. Labels which may be employed as part of a signal producing system capable of producing a signal in relation to the amount of analyte in the sample include free radicals, chromogens, such as fluorescent dyes, 5 chemiluminescers, enzymes, bacteriophages, coenzymes particulate labels and so forth.
- In a heterogeneous assay approach, the reagents are usually the sample, the specific antibody, and means for producing a detectable signal. The specimen is generally placed on a support, such as a plate or a slide, and contacted with the antibody in a liquid phase. The support is then separated from the liquid phase and either the support phase or the liquid phase is examined 10 for a detectable signal employing means for producing such signal or signal producing system. The signal is related to the presence of the analyte in the sample. Means for producing a detectable signal includes the use of radioactive labels, fluorescers, enzymes, and so forth. Exemplary of heterogeneous immunoassays are the radioimmunoassay, immunofluorescence methods, enzyme-linked immunoassays, and the like.
- 15 One embodiment of an assay employing an antibody of the present invention involves the use of a surface to which the monoclonal antibody of the invention is attached. The underlying structure of the surface may take different forms, have different compositions and may be a mixture of compositions or laminates or comthereof. The surface may assume a variety of shapes and -forms and may have varied dimensions, depending on the manner of use and measurement.
- 20 Illustrative surfaces may be pads, beads, discs, or strips which may be flat, concave or convex. Thickness is not critical, generally being from about 0.1 to 2 mm thick and of any convenient diameter or other dimensions. surface typically will be supported on a rod, tube, capillary, fiber, strip, disc, plate, cuvette and the like. The surface will typically be porous and polyfunctional or capable of being polyfunctionalized so as to permit covalent binding of the monoclonal antibody of 25 the invention as well as to permit bonding of other compounds which form a part of a means for producing a detectable signal. A wide variety of organic and inorganic polymers, both natural and synthetic, and combinations thereof, may be employed as the material for the solid surface. Illustrative polymers include polyethylene, polypropylene, poly(4-methylbutene), polystyrene, polymethacrylate, poly(ethylene terephthalate), rayon, nylon, poly(vinyl butyrate), silicones,
- 30 polyformaldehyde, cellulose, cellulose acetate, nitrocellulose, latex, etc. Other which may be employed include paper, glasses, ceramics, metals, metalloids, semiconductor materials, cermets, silicates or the like. Also included substrates that form gels, gelatins, lipopolysaccharides, silicates, agarose and polyacrylamides or polymers which form several aqueous phases such as dextrans, polyalkylene glycols (alkylene of 2 to 3 carbon atoms) or surfactants such as phospholipids. The 35 binding of the monoclonal antibody of the invention to the surface may be accomplished by well known techniques, commonly available in the literature. See, for example, "Immobilized Enzymes," Ichiro Chibata, Press, New York (1978) and Cuatrecasas, J Bio. Chem., 245: 3059

(1970). In carrying out the assay in accordance with this of the invention the sample is mixed with aqueous medium and the medium is contacted with the surface having a monoclonal antibody of the invention bound thereto. Members of a signal producing system and any ancillary materials may also be included in the aqueous medium, either concurrently or added subsequently so as to provide a detectable signal associated with the surface. The means for producing the detectable signal can involve the incorporation in the of a labeled analytic or it may involve the use of a second monoclonal antibody having a label conjugated thereto. Separation and washing steps will be carried out as needed. The signal detected is related to the presence of Pro412-related proteins in the sample. It is within the of the present invention to include a calibration as the measurement surface on the same support. A particular embodiment of an assay in accordance with the present invention, by way of illustration and not limitation, involves the use of a support such as a slide or a well of a petri dish. The technique involves the sample to be analyzed on the support with an appropriate fixing material such as acetone and incubating the sample on the slide with a monoclonal antibody of the invention. After washing with an appropriate buffer such as, for example, phosphate buffered saline, the support is contacted with a labeled specific binding partner for the analyte in the sample. After incubation as desired, the slide is washed a second time with an aqueous buffer and the determination is made of the binding of the labeled monoclonal antibody to the analytic. If the label is fluorescent, the slide may be covered with a fluorescent antibody mounting fluid on a cover slip and then examined with a fluorescent microscope to determine the extent of binding. On the other hand, the label can be an enzyme conjugated to the monoclonal antibody of the invention and the extent of binding can be determined by examining the slide for the presence of enzyme activity, which may be indicated by the formation of a precipitate, a color, or the like. A particular example of an assay utilizing the present antibodies is a double determinant ELISA assay. A support such as, e.g., a glass or vinyl plate, is coated with antibody specific for Pro412-related proteins by conventional techniques. The support is contacted with the sample suspected of containing Pro412-related proteins, usually in a aqueous medium. After an incubation period from 30 seconds to 12 hours, the support is separated from the medium, washed to remove unbound Pro412-related proteins with, for example, water or an aqueous buffered medium, and contacted with an antibody specific for Pro412-related proteins, again usually in an aqueous medium. The antibody is labeled with an enzyme directly or indirectly such as, e.g., horseradish peroxidase or alkaline phosphatase. After incubation, the support is separated from the medium, and washed as above. The enzyme activity of the support or the aqueous medium is determined. This enzyme activity is related to the amount of Pro412-related proteins in the sample.

The invention also includes kits, e.g., diagnostic assay kits, for carrying out the methods disclosed above. In one embodiment, the kit comprises in packaged combination (a) a monoclonal antibody more specifically defined above and (b) a conjugate of a specific binding partner for the

- above monoclonal antibody and a label capable of producing a detectable signal. The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The kit may further include, where necessary, other members of the signal producing system of which system the label is a member, agents for reducing background interference in a test, control reagents, apparatus for conducting a test, and the like. In another embodiment, the diagnostic kit comprises a conjugate of monoclonal antibody of the invention and a label capable of producing a detectable signal. Ancillary agents as mentioned above may also be present.
- 5 interference in a test, control reagents, apparatus for conducting a test, and the like. In another embodiment, the diagnostic kit comprises a conjugate of monoclonal antibody of the invention and a label capable of producing a detectable signal. Ancillary agents as mentioned above may also be present.
- 10 An anti-Pro412-related antibody (e.g., monoclonal antibody) can be used to isolate Pro412-related protein by standard techniques, such as affinity chromatography or immunoprecipitation. For example, an anti-Pro412-related antibody can facilitate the purification of natural Pro412-related proteins from cells and of recombinantly produced Pro412-related proteins expressed in host cells. Moreover, an anti-Pro412-related antibody can be used to detect Pro412-related protein
- 15 (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the Pro412-related protein. Anti-Pro412-related antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable
- 20 substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliflourone, fluorescein, fluorescein
- 25 isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .
- 30 *Drug screening assays*
- The invention provides a method (also referred to herein as a "screening assay") for identifying inhibitors, i.e., candidate or test compounds or agents (e.g., small molecules and peptides, antibodies, peptidomimetics or other drugs) which bind to Pro412-related proteins, have an inhibitory or activating effect on, for example, Pro412-related proteins expression or preferably
- 35 Pro412-related activity, or have an inhibitory or activating effect on, for example, the activity of a Pro412-related target molecule. In some embodiments small molecules can be generated using combinatorial chemistry or can be obtained from a natural products library. Assays may be cell

based or non-cell based assays. Drug screening assays may be binding assays or more preferentially functional assays, as further described.

In other embodiments, an assay is a cell-based assay in which a cell which expresses a Pro412-related protein or biologically active portion thereof is contacted with a test compound and 5 the ability of the test compound to inhibit, activate, or increase Pro412-related activity determined. Determining the ability of the test compound to inhibit, activate, or increase Pro412-related activity can be accomplished by monitoring the bioactivity of the Pro412-related protein or biologically active portion thereof. The cell, for example, can be of mammalian origin, insect origin, bacterial origin or a yeast cell. For example, in some embodiments, the cell can be a 10 mammalian cell, bacterial cell or yeast cell which has been engineered to lack a natural inhibitor of a Pro412-related proteins.

The invention further encompasses compounds capable of inhibiting or activating Pro412-related activity. Preferably, a Pro412-related inhibitor or activator is a selective Pro412-related 15 inhibitor or activator. Assays of the invention may be used to screen any suitable collection of compounds.

In one embodiment, the invention provides assays for screening candidate or test compounds which are target molecules of a Pro412-related protein or polypeptide or biologically 20 active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a Pro412-related protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase 25 libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is used with peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K. S. (1997) Anticancer Drug Des. 12:145, the disclosure of which is incorporated herein by reference in its 30 entirety).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) Proc. Natl. Acad. Sci. U.S.A. 90:6909; Erb et al. (1994) Proc. Natl. Acad. Sci. USA 91:11422; Zuckermann et al. (1994). J. Med. Chem. 37:2678; Cho et al. (1993) Science 261:1303; Carrell et al. (1994) Angew. Chem. Int. Ed. Engl. 33:2059; Carell et al. 35 (1994) Angew. Chem. Int. Ed. Engl. 33:2061; and in Gallop et al (1994) J. Med. Chem. 37:1233, the disclosures of which are incorporated herein by reference in their entireties.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) Biotechniques 13:412-421), or on beads (Lam (1991) Nature 354:82-84), chips (Fodor (1993)

Nature 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner U.S. Pat. No. '409), plasmids (Cull et al. (1992) Proc Natl Acad Sci USA 89:1865-1869) or on phage (Scott and Smith (1990) Science 249:386-390); (Devin (1990) Science 249:404-406); (Cwirla et al. (1990) Proc. Natl. Acad. Sci. 87:6378-6382); (Felici (1991) J. Mol. Biol. 222:301-310); (Ladner *supra*),
5 the disclosures of which are incorporated herein by reference in their entireties.

Determining the ability of the test compound to inhibit or increase Pro412-related activity can also be accomplished, for example, by coupling the Pro412-related protein or biologically active portion thereof with a radioisotope or enzymatic label such that binding of the Pro412-
10 related protein or biologically active portion thereof to its cognate target molecule can be determined by detecting the labeled Pro412-related protein or biologically active portion thereof in a complex. For example, compounds (e.g., Pro412-related protein or biologically active portion thereof) can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioimmersion or by scintillation counting
15 Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. The labeled molecule is placed in contact with its cognate molecule and the extent of complex formation is measured. For example, the extent of complex formation may be measured by immuno precipitating the complex or by performing gel
20 electrophoresis.

It is also within the scope of this invention to determine the ability of a compound (e.g., Pro412-related protein or biologically active portion thereof) to interact with its cognate target molecule without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with its cognate target molecule without the labeling
25 of either the compound or the target molecule. McConnell, H. M. et al. (1992) Science 257:1906-1912, the disclosure of which is incorporated herein by reference in its entirety. A microphysiometer such as a cytosensor is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between compound and
30 receptor.

In a preferred embodiment, the assay comprises contacting a cell which expresses a Pro412-related protein or biologically active portion thereof, with a target molecule to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to inhibit or increase the activity of the Pro412-related protein or biologically active portion thereof, wherein determining the ability of the test compound to inhibit or increase the activity of the Pro412-related protein or biologically active portion thereof, comprises
35 determining the ability of the test compound to inhibit or increase a biological activity of the

Pro412-related proteins expressing cell (e.g., determining the ability of the test compound to inhibit or increase transduction, protein:protein interactions, proteolysis (as discussed above).

In another preferred embodiment, the assay comprises contacting a cell which is responsive to a Pro412-related protein or biologically active portion thereof, with a Pro412-related

- 5 protein or biologically-active portion thereof, to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to modulate the activity of the Pro412-related protein or biologically active portion thereof, wherein determining the ability of the test compound to modulate the activity of the Pro412-related protein or biologically active portion thereof comprises determining the ability of the test compound to
- 10 modulate a biological activity of the Pro412-related proteins-responsive cell (e.g., determining the ability of the test compound to modulate an indirect Pro412-related activity such as e.g. signal transduction or protein:protein interactions, or determining the ability of the test compound to modulate antigen processing, cell proliferation or cell death).

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a Pro412-related target molecule (i.e. a molecule with which Pro412-related proteins interact, preferably a molecule cleaved by Pro412-related proteins) with a test compound and determining the ability of the test compound to modulate (e.g. stimulate or inhibit) the activity of the Pro412-related target molecule. Determining the ability of the test compound to modulate the activity of a Pro412-related target molecule can be accomplished, for example, by determining assessing the activity of a target molecule (that is, of a cleavage substrate or of the cleavage product), or by assessing the ability of the Pro412-related protein to bind to or interact with the Pro412-related target molecule.

Determining the ability of the Pro412-related protein to bind to or interact with a Pro412-related target molecule, such as for example a natural Pro412-related inhibitor, can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the Pro412-related protein to bind to or interact with a Pro412-related target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by contacting the target molecule with the Pro412-related protein or a fragment thereof and measuring induction of a cellular second messenger of the target (i.e. intracellular Ca²⁺, diacylglycerol, IP₃, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular response, for example, signal transduction or protein:protein interactions

35 In yet another embodiment, an assay of the present invention is a cell-free assay in which a Pro412-related protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the Pro412-related protein or biologically active portion

- thereof is determined. Binding of the test compound to the Pro412-related protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the Pro412-related protein or biologically active portion thereof with a known compound which binds Pro412-related proteins (e.g., a Pro412-related target molecule) to form an
- 5 assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a Pro412-related protein, wherein determining the ability of the test compound to interact with a Pro412-related protein comprises determining the ability of the test compound to preferentially bind to Pro412-related proteins or biologically active portion thereof as compared to the known compound.
- 10 In another embodiment, the assay is a cell-free assay in which a Pro412-related protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the Pro412-related protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a Pro412-related protein can be accomplished, for example, by
- 15 determining the ability of the Pro412-related protein to bind to a Pro412-related target molecule by one of the methods described above for determining direct binding. Determining the ability of the Pro412-related protein to bind to a Pro412-related target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) Anal. Chem. 63:2338-2345 and Szabo et al. (1995) Curr. Opin. Struct. Biol. 5:699-705, the disclosures of which are incorporated herein by reference in their entireties. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIACore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.
- 20 25 In an alternative embodiment, determining the ability of the test compound to modulate the activity of a Pro412-related protein can be accomplished by determining the ability of the Pro412-related protein to further modulate the activity of a downstream effector (e.g., a growth factor mediated signal transduction pathway component) of a Pro412-related target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.
- In yet another embodiment, the cell-free assay involves contacting a Pro412-related protein or biologically active portion thereof with a known compound which binds the Pro412-related protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the Pro412-related protein, wherein
- 30 35 determining the ability of the test compound to interact with the Pro412-related protein comprises determining the ability of the Pro412-related protein to preferentially bind to or modulate the activity of a Pro412-related target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (e.g. Pro412-related proteins or biologically active portions thereof or molecules to which Pro412-related targets bind). In the case of cell-free assays in which a membrane-bound form an isolated protein is used it may be desirable to utilize a

- 5 solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton TM X-100, Triton TM X-114, Thesit TM, Isotridecypoly(ethylene glycol ether)n,3-[{3-cholamidopropyl}dimethylammonio]-1-propane sulfonate (CHAPS), 3-[{3-cholamidopropyl}dimethylammonio]-2-hydroxy-1-propan sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.
- 10

- 15 In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either a Pro412-related protein or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a Pro412-related protein, or interaction of a Pro412-related protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the
- 20 proteins to be bound to a matrix. For example, glutathione-S-transferase/Pro412-related fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or Pro412-related protein, and the mixture incubated under conditions conducive to
- 25 complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of Pro412-related proteins binding or activity determined using standard techniques.

- 30 Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a Pro412-related protein or a Pro412-related target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated Pro412-related protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical).
- 35 Alternatively, antibodies reactive with Pro412-related protein or target molecules but which do not interfere with binding of the Pro412-related protein to its target molecule can be derivatized to the

wells of the plate, and unbound target or Pro412-related protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the Pro412-related protein or target molecule, as well as enzyme-linked assays which rely on 5 detecting an enzymatic activity associated with the Pro412-related protein or target molecule.

In another embodiment, modulators of Pro412-related proteins expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of Pro412-related proteins mRNA or protein in the cell is determined. The level of expression of Pro412-related proteins mRNA or protein in the presence of the candidate compound is compared to the 10 level of expression of Pro412-related proteins mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of Pro412-related proteins expression based on this comparison. For example, when expression of Pro412-related proteins mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of 15 Pro412-related proteins mRNA or protein expression. Alternatively, when expression of Pro412-related proteins mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of Pro412-related proteins mRNA or protein expression. The level of Pro412-related proteins mRNA or protein expression in the cells can be determined by methods described herein for detecting 20 Pro412-related proteins mRNA or protein.

In yet another aspect of the invention, the Pro412-related proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J. Biol. Chem. 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696; and Brent 25 WO94/10300, the disclosures of which are incorporated herein by reference in their entireties), to identify other proteins, which bind to or interact with Pro412-related proteins ("Pro412-related proteins-binding proteins" or "Pro412-related proteins-bp") and are involved in Pro412-related activity. Such Pro412-related proteins-binding proteins are also likely to be involved in the propagation of signals by the Pro412-related proteins or Pro412-related targets as, for example, 30 downstream elements of a Pro412-related proteins-mediated signaling pathway. Alternatively, such Pro412-related proteins-binding proteins are likely to be Pro412-related inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a Pro412-related protein or a fragment thereof is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prcy" or "sample") is fused to a gene that codes for the activation domain 35

of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a Pro412-related proteins-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to
5 the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the Pro412-related protein.

This invention further pertains to novel agents identified by the above-described screening assays and to processes for producing such agents by use of these assays. Accordingly, in one embodiment, the present invention includes a compound or agent obtainable by a method comprising the steps of any one of the aforementioned screening assays (e.g., cell-based assays or cell-free assays). For example, in one embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a Pro412-related target molecule with a test compound and determining the ability of the test compound to bind to, or modulate the activity of, the Pro412-related target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a Pro412-related target molecule with a Pro412-related protein or biologically-active portion thereof, to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of, the
10 Pro412-related target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a Pro412-related protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to, or modulate (e.g., stimulate or inhibit) the activity of, the Pro412-related protein or biologically active portion thereof. In yet another embodiment, the present invention included a compound or
15 agent obtainable by a method comprising contacting a Pro412-related protein or biologically active portion thereof with a known compound which binds the Pro412-related protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of the Pro412-related protein.
20

Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a Pro412-related proteins modulating agent, an antisense Pro412-related nucleic acid molecule, a Pro412-related proteins-specific antibody, or a Pro412-related proteins-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.
25

The present invention also pertains to uses of novel agents identified by the above-described screening assays for diagnoses, prognoses, and treatments as described herein.

Accordingly, it is within the scope of the present invention to use such agents in the design,

- 5 formulation, synthesis, manufacture, and/or production of a drug or pharmaceutical composition for use in diagnosis, prognosis, or treatment, as described herein. For example, in one embodiment, the present invention includes a method of synthesizing or producing a drug or pharmaceutical composition by reference to the structure and/or properties of a compound obtainable by one of the above-described screening assays. For example, a drug or pharmaceutical composition can be
- 10 synthesized based on the structure and/or properties of a compound obtained by a method in which a cell which expresses a Pro412-related target molecule is contacted with a test compound and the ability of the test compound to bind to, or modulate the activity of, the Pro412-related target molecule is determined. In another exemplary embodiment, the present invention includes a method of synthesizing or producing a drug or pharmaceutical composition based on the structure
- 15 and/or properties of a compound obtainable by a method in which a Pro412-related protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to, or modulate (e.g., stimulate or inhibit) the activity of, the Pro412-related protein or biologically active portion thereof is determined.

20 *Pharmaceutical Compositions*

When polypeptides of the present invention are expressed in soluble form, for example as a secreted product of transformed yeast or mammalian cells, they can be purified according to standard procedures of the art, including steps of ammonium sulfate precipitation, ion exchange chromatography, gel filtration, electrophoresis, affinity chromatography, and/or the like, e.g.

- 25 "Enzyme Purification and Related Techniques," Methods in Enzymology, 22:233-577 (1977), and Scopes, R., Protein Purification: Principles and Practice (Springer-Verlag, New York, 1982) provide guidance in such purifications. Likewise, when polypeptides of the invention are expressed in insoluble form, for example as aggregates, inclusion bodies, or the like, they can be purified by standard procedures in the art, including separating the inclusion bodies from disrupted
- 30 host cells by centrifugation, solubilizing the inclusion bodies with chaotropic and reducing agents, diluting the solubilized mixture, and lowering the concentration of chaotropic agent and reducing agent so that the polypeptide takes on a biologically active conformation. The latter procedures are disclosed in the following references, which are incorporated by reference. Winkler et al, Biochemistry, 25: 4041-4045 (1986), Winkler et al, Biotechnology, 3: 992-998 (1985); Koths et al, U.S. patent 4,569,790; and European patent applications 86306917.5 and 86306353.3.

As used herein "effective amount" means an amount sufficient to ameliorate a symptom of an autoimmune condition. The effective amount for a particular patient may vary depending on such factors as the state of the condition being treated, the overall health of the patient, method of

administration, the severity of side-effects, and the like.

- Pro412-related proteins can be purified from culture supernatants of mammalian cells transiently transfected or stably transformed by an expression vector carrying a Pro412-related gene. Preferably, Pro412-related proteins is purified from culture supernatants of COS 7 cells
- 5 transiently transfected by the pcD expression vector. Transfection of COS 7 cells with pcD proceeds as follows: One day prior to transfection, approximately 10^6 COS 7 monkey cells are seeded onto individual 100 mm plates in Dulbecco's modified Eagle medium (DME) containing 10% fetal calf serum and 2 mM glutamine. To perform the transfection, the medium is aspirated from each plate and replaced with 4 ml of DME containing 50 mM Tris.HCl pH 7.4, 400 mg/ml
- 10 DEAE-Dextran and 50 µg of plasmid DNA. The plates are incubated for four hours at 37°C, then the DNA-containing medium is removed, and the plates are washed twice with 5 ml of serum-free DME. DME is added back to the plates which are then incubated for an additional 3 hrs at 37°C. The plates are washed once with DME, after which DME containing 4% fetal calf serum, 2 mM glutamine, penicillin (100 U/L) and streptomycin (100 µg/L) at standard concentrations is added.
- 15 The cells are then incubated for 72 hrs at 37°C, after which the growth medium is collected for purification of Pro412-related proteins. Alternatively, transfection can be accomplished by electroporation as described in the examples. Plasmid DNA for the transfections is obtained by growing pcD(SR α), or like expression vector, containing the Pro412-related proteins cDNA insert in E. coli MC1061, described by Casadaban and Cohen, J Mol. Biol., Vol. 138, pgs. 179-207
- 20 (1980), or like organism. The plasmid DNA is isolated from the cultures by standard techniques, e.g. Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor Laboratory, New York, 1989) or Ausubel et al (1990, cited above).

- Compounds capable of inhibiting a Pro412-related activity, preferably small molecules but also including peptides, Pro412-related nucleic acid molecules, Pro412-related proteins, and anti-Pro412-related antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.
- 35 A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for

injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as 5 sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic

- Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of 10 sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL® (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi.
- 15 The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action microorganisms can be achieved by various 20 antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the 25 composition an agent which delays absorption, for example, aluminum monostearate and gelatin.
- Where the active compound is a protein, peptide or anti-Pro412-related antibody, sterile injectable solutions can be prepared by incorporating the active compound (e.g.,) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the 30 required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.
- Oral compositions generally include an inert diluent or an edible carrier. They can be 35 enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. For administration by inhalation, the compounds are delivered in the

form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer. Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Most preferably, active compound is delivered to a subject by intravenous injection.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811, the disclosure of which is incorporated herein by reference in its entirety.

It is especially advantageous to formulate oral or preferably parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or NO toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

15 *Methods of Treatment*

Provided herein are treatment methods based on the Pro412-related proteins of the invention. Both active Pro412-related proteins as provided in SEQ ID NOS 1 to 4, as well as inhibitors of Pro412-related proteins such as antibodies may be useful in the treatment or prevention of disease.

20 Modulation (increasing or decreasing) of Pro412-related activity can be carried out in a number of suitable ways, several of which have been described in the present application. For example, methods of treatment may involve modulating a Pro412-related activity, biological activity of a Pro412-related protein or functional activity of a Pro412-related protein. Modulating Pro412-related activity may involve modulating an association with a Pro412-related-target molecule (e.g. a substrate) or preferably any other Pro412-related activity as discussed herein. Preferably modulating a Pro412-related activity comprises modulating proteolytic activity.

30 In one aspect, the Pro412-related proteins of the invention may be used for stimulating endothelial cell growth in a mammal, particularly for the stimulation of angiogenesis. In another aspect, a Pro412-related protein may be useful in diagnosing a cardiovascular, angiogenic or endothelial disorder in a mammal which comprises analyzing the level of expression of a nucleic acids encoding the Pro412-related proteins of the invention, or analyzing the level of Pro412-related proteins of the invention.

35 In one embodiment, vascularization of tumors is attacked using antibodies that bind specifically to the Pro412-related proteins of the invention. The anti-Pro412-related antibody and another are

administered to tumor-bearing patients at therapeutically effective doses as determined, for example, by observing necrosis of the tumor or its metastatic foci, if any. This therapy is continued until such time as no further beneficial effect is observed or clinical examination shows no trace of the tumor or any metastatic foci.

- 5 In other aspect, the Pro412-related protein or agonists or antagonists thereto may also be employed for their cell growth stimulating activity to stimulate erythropoiesis or granulopoiesis, to stimulate wound healing or tissue regeneration and associated therapies concerned with re-growth of tissue, such as connective tissue, skin, bone, cartilage, muscle, liver or kidney, to promote angiogenesis, to stimulate growth of endothelial cells, and to proliferate the growth of vascular smooth muscle and endothelial cell production. The increase in angiogenesis mediated by the Pro412-related protein or antagonist would be beneficial for the treatment of ischemic tissues and to collateral coronary development in the heart subsequent to coronary stenosis.
- 10 A Pro412-related protein can be used for its angiogenic activity in peripheral vascular disease, hypertension. Inflammatory vasculitides, Reynaud's disease and Reynaud's phenomenon, aneurysms, arterial restenosis, wound healing and tissue repair, ischemia reperfusion injury, angina, myocardial infarctions such as acute myocardial infarctions, chronic heart conditions, heart failure such as congestive heart failure, and osteoporosis. Hypertension for example is characterized by raised vascular pressure in the systemic arterial, pulmonary arterial, or portalvenous systems. Elevated pressure may result from or result in impaired endothelial function
- 15 and/or vascular disease. In other aspects, antibodies of antagonists to Pro412-related proteins can be used for inhibiting endothelial cell growth, and thus for the treatment of chronic inflammatory disorders such as rheumatoid arthritis or Crohn's disease

- 20 Pro412-related proteins can also be used for stimulating the activity of T-cells, preferably in methods comprising contacting CD4+ cells or PBMC cells with an effective amount of a Pro412-related protein. In preferred method, the invention comprises administering to the mammal an effective amount of a Pro412-related protein of the invention for the treatment of a tumor. In other aspects an antagonist of a Pro412-related protein is used for the treatment of an the immune related disorder is selected from the group consisting of systemic lupus erythematosus, rheumatoid
- 25 arthritis, osteoarthritis, juvenile chronic arthritis, idiopathic inflammatory myopathies, Sjogren's syndrome, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central and peripheral nervous systems such as multiple sclerosis, inflammatory bowel disease, psoriasis, allergic diseases such as asthma, allergic rhinitis, atopic dermatitis, and transplantation associated diseases including graft rejection and graft -
- 30 versus-host-disease.

In yet further aspects, the Pro412-related proteins of the invention may be used for promoting the

survival of retinal cells. Thus, encompassed are methods of delaying, preventing or rescuing photoreceptor cells from injury or death in a mammal comprising administrating a therapeutically effective amount of a Pro412-related protein to said mammal. Pro412-related protein can be used to delay, prevent or rescue retinal cells (including retinal neurons; e.g., photoreceptor cells, other 5 retinal neurons, and supportive cells from injury or death. Other retinal neurons include, but are not limited to retinal ganglion cells, displaced retinal ganglion cells, amacrine cells, displaced amacrine cells, horizontal and bipolar neurons. Supportive cells include, e.g., Muller cells or RPE cells. Pro412-related proteins to treat any condition which results in injury or death or 10 photoreceptor cells or other retinal cells. Examples of such conditions include: retinal detachment; age-related and other maculopathies; photic retinopathies, surgery-induced retinopathies (either mechanically or light-induced), toxic retinopathies including those resulting from foreign bodies in the eye; diabetic retinopathies; retinopathy of prematurity; viral retinopathies such as CMV or HIV 15 retinopathy related to AIDS; uveitis; ischemic retinopathies due to venous or arterial occlusion or other vascular disorder; retinopathies due to trauma or penetrating lesions of the eye; peripheral vitreoretinopathy; and inherited retinal degenerations. Exemplary retinal degenerations include e.g., hereditary spastic paraparesis with retinal degeneration (Kjellin and Barnard-Scholtz 20 syndromes), retinitis pigmentosa, Stargardt disease, Usher syndrome (retinitis pigmentosa with congenital hearing loss), and Refsum syndrome (retinitis pigmentosa, hereditary hearing loss and polyneuropathy). Additional disorders which result in death or retinal neurons include, retinal tears, detachment of the retina and pigment epithelium, degenerative myopia, acute retinal necrosis syndrome (ARN), traumatic chorioretinopathies or contusion (Purtscher's Retinopathy) and edema.

In view of their role in stimulating growth of skin cells, Pro412-related proteins can also be used in methods for treating skin disorders. Examples of skin disorders include skin cancers (e.g., 25 melanomas), eczema, and psoriasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g., skin and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, 30 urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level.

Diagnostic and Prognostic Uses

35 The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics; and in drug screening and methods of treatment

(e.g., therapeutic and prophylactic) as further described herein.

The invention provides diagnostic and prognostic assays for detecting Pro412-related nucleic acids and proteins, as further described. Also provided are diagnostic and prognostic assays for detecting interactions between Pro412-related proteins and Pro412-related target molecules, particularly natural inhibitors or cleavage substrates.

The isolated nucleic acid molecules of the invention can be used, for example, to detect Pro412-related proteins mRNA (e.g., in a biological sample) or a genetic alteration in a Pro412-related gene, and to modulate a Pro412-related activity, as described further below. The Pro412-related proteins can be used to treat disorders characterized by insufficient or excessive production of a Pro412-related protein or Pro412-related target molecule. In addition, the Pro412-related proteins can be used to screen for naturally occurring Pro412-related target molecules, to screen for drugs or compounds which modulate, preferably inhibit Pro412-related activity, as well as to treat disorders characterized by insufficient or excessive production of Pro412-related protein or production of Pro412-related protein forms which have decreased or aberrant activity compared to Pro412-related proteins wild type protein. Moreover, the anti-Pro412-related antibodies of the invention can be used to detect and isolate Pro412-related proteins, regulate the bioavailability of Pro412-related proteins, and modulate Pro412-related activity.

Accordingly one embodiment of the present invention involves a method of use (e.g., a diagnostic assay, prognostic assay, or a prophylactic/therapeutic method of treatment) wherein a molecule of the present invention (e.g., a Pro412-related protein, Pro412-related nucleic acid, or most preferably a Pro412-related inhibitor or activator) is used, for example, to diagnose, prognosis and/or treat a disease and/or condition in which any of the aforementioned Pro412-related proteins activities is indicated. In another embodiment, the present invention involves a method of use (e.g., a diagnostic assay, prognostic assay, or a prophylactic/therapeutic method of treatment) wherein a molecule of the present invention (e.g., a Pro412-related protein, Pro412-related nucleic acid, or a Pro412-related inhibitor or activator) is used, for example, for the diagnosis, prognosis, and/or treatment of subjects, preferably a human subject, in which any of the aforementioned activities is pathologically perturbed. In a preferred embodiment, the methods of use (e.g., diagnostic assays, prognostic assays, or prophylactic/therapeutic methods of treatment) involve administering to a subject, preferably a human subject, a molecule of the present invention (e.g., a Pro412-related protein, Pro412-related nucleic acid, or a Pro412-related inhibitor or activator) for the diagnosis, prognosis, and/or therapeutic treatment. In another embodiment, the methods of use (e.g., diagnostic assays, prognostic assays, or prophylactic/therapeutic methods of treatment) involve administering to a human subject a molecule of the present invention (e.g., a Pro412-related protein, Pro412-related nucleic acid, or a Pro412-related inhibitor or activator).

For example, the invention encompasses a method of determining whether a Pro412-related proteins member is expressed within a biological sample comprising: a) contacting said

- biological sample with: ii) a polynucleotide that hybridizes under stringent conditions to a Pro412-related nucleic acid; or iii) a detectable polypeptide (e.g. antibody) that selectively binds to a Pro412-related protein; and b) detecting the presence or absence of hybridization between said polynucleotide and an RNA species within said sample, or the presence or absence of binding of 5 said detectable polypeptide to a polypeptide within said sample. A detection of said hybridization or of said binding indicates that said Pro412-related proteins member is expressed within said sample. Preferably, the polynucleotide is a primer, and wherein said hybridization is detected by detecting the presence of an amplification product comprising said primer sequence, or the detectable polypeptide is an antibody.
- 10 Also envisioned is a method of determining whether a mammal, preferably human, has an elevated or reduced level of expression of a Pro412-related proteins member, comprising: a) providing a biological sample from said mammal; and b) comparing the amount of a Pro412-related protein or of a Pro412-related proteins RNA species encoding a Pro412-related protein within said biological sample with a level detected in or expected from a control sample. An 15 increased amount of said Pro412-related protein or said Pro412-related proteins RNA species within said biological sample compared to said level detected in or expected from said control sample indicates that said mammal has an elevated level of Pro412-related proteins expression, and a decreased amount of said Pro412-related protein or said Pro412-related proteins RNA species within said biological sample compared to said level detected in or expected from said control 20 sample indicates that said mammal has a reduced level of expression of a Pro412-related proteins member.

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present 25 invention relates to diagnostic assays for determining Pro412-related protein and/or nucleic acid expression as well as Pro412-related activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant Pro412-related proteins expression or activity. The invention also provides for prognostic (or predictive) assays for 30 determining whether an individual is at risk of developing a disorder associated with a Pro412-related protein, nucleic acid expression or activity. For example, mutations in a Pro412-related gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with a Pro412-related protein, nucleic acid expression or activity.

35 Accordingly, the methods of the present invention are applicable generally to diseases related to regulation of proteolysis, including but not limited to disorders characterized by unwanted cell proliferation or generally aberrant control of differentiation, for example neoplastic

or hyperplastic disorders, as well as disorders related to proliferation or lack thereof of endothelial cells, inflammatory disorders and neurodegenerative disorders.

Diagnostic Assays

5

An exemplary method for detecting the presence (quantitative or not) or absence of a Pro412-related protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting a Pro412-related protein or nucleic acid (e.g., mRNA, Pro412-related proteins 10 genomic DNA) that encodes Pro412-related protein such that the presence of the Pro412-related protein or nucleic acid is detected in the biological sample. A preferred agent for detecting a Pro412-related protein mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to a Pro412-related protein mRNA or Pro412-related protein genomic DNA. The nucleic acid probe can be, for example, a full-length Pro412-related nucleic acid, or a probe 15 sufficient in length to specifically hybridize under stringent conditions to a Pro412-related protein mRNA or genomic DNA or a portion of a Pro412-related nucleic acid. Other suitable probes for use in the diagnostic assays of the invention are described herein.

In preferred embodiments, the subject method can be characterized by generally comprising detecting, in a tissue sample of the subject (e.g. a human patient), the presence or 20 absence of a genetic lesion characterized by at least one of (i) a mutation of a gene encoding one of the subject Pro412-related proteins or (ii) the mis-expression of a Pro412-related gene. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from a Pro412-related gene, (ii) an addition of one or more nucleotides to such a Pro412-related gene, (iii) a substitution of one or more nucleotides of a 25 Pro412-related gene, (iv) a gross chromosomal rearrangement or amplification of a Pro412-related gene, (v) a gross alteration in the level of a messenger RNA transcript of a Pro412-related gene, (vi) aberrant modification of a Pro412-related gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a Pro412-related gene, and (viii) a non-wild type level of a Pro412-related target 30 protein.

A preferred agent for detecting a Pro412-related protein is an antibody capable of selectively binding to a Pro412-related protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, 35 is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling

- include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the
- 5 detection method of the invention can be used to detect a Pro412-related proteins mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of a Pro412-related proteins mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a Pro412-related protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence.
- 10 *In vitro* techniques for detection of a Pro412-related proteins genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a Pro412-related protein include introducing into a subject a labeled anti-Pro412-related antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.
- 15 In yet another exemplary embodiment, aberrant methylation patterns of a Pro412-related nucleic acid can be detected by digesting genomic DNA from a patient sample with one or more restriction endonucleases that are sensitive to methylation and for which recognition sites exist in the Pro412-related gene (including in the flanking and intronic sequences). See, for example, Buiting et al. (1994) Human Mol Genet 3:893-895. Digested DNA is separated by gel
- 20 electrophoresis, and hybridized with probes derived from, for example, genomic or cDNA sequences. The methylation status of the Pro412-related gene can be determined by comparison of the restriction pattern generated from the sample DNA with that for a standard of known methylation.
- 25 Furthermore, gene constructs such as those described herein can be utilized in diagnostic assays to determine if a cell's growth or differentiation state is no longer dependent on the function of a Pro412-related protein, e.g. in determining the phenotype of a transformed cell. Such knowledge can have both prognostic and therapeutic benefits. To illustrate, a sample of cells from the tissue can be obtained from a patient and dispersed in appropriate cell culture media, a portion of the cells in the sample can be caused to express a recombinant Pro412-related protein or a
- 30 Pro412-related target protein, e.g. by transfection with a expression vector described herein, or to increase the expression or activity of an endogenous Pro412-related protein or Pro412-related target protein, and subsequent growth of the cells assessed. The absence of a change in phenotype of the cells despite expression of the Pro412-related proteins or Pro412-related target protein may be indicative of a lack of dependence on cell regulatory pathways which includes the Pro412-
- 35 related proteins or Pro412-related target protein, e.g. Pro412-related proteins- or Pro412-related target-mediated transcription. Depending on the nature of the tissue of interest, the sample can be in the form of cells isolated from, for example, a blood sample, an exfoliated cell sample, a fine

needle aspirant sample, or a biopsied tissue sample. Where the initial sample is a solid mass, the tissue sample can be minced or otherwise dispersed so that cells can be cultured, as is known in the art.

- In yet another embodiment, a diagnostic assay is provided which detects the ability of a
- 5 Pro412-related gene product, e.g., isolated from a biopsied tissue, to bind to other cellular, cell surface or extracellular proteins. For instance, it will be desirable to detect Pro412-related protein mutants which, while expressed at appreciable levels in the cell, are defective at binding or preferably cleaving a Pro412-related target protein (having either diminished or enhanced proteolytic activity or binding affinity for the target). Such mutants may arise, for example, from
- 10 mutations, e.g., point mutants, which may be impractical to detect by the diagnostic DNA sequencing techniques or by the immunoassays described above. The present invention accordingly further contemplates diagnostic screening assays which generally comprise cloning one or more Pro412-related genes from the sample tissue, and expressing the cloned genes under conditions which permit detection of an interaction between that recombinant gene product and a
- 15 target protein, or under conditions which permit assessment of proteolytic activity. As will be apparent from the description of the various drug screening assays set forth herein, a wide variety of techniques can be used to determine the ability of a Pro412-related protein to bind to other components. These techniques can be used to detect mutations in a Pro412-related gene which give rise to mutant proteins with a higher or lower binding affinity for a Pro412-related target protein relative to the wild-type Pro412-related protein. Conversely, by switching which of the Pro412-related target protein and Pro412-related protein is the "bait" and which is derived from the patient sample, the subject assay can also be used to detect Pro412-related target protein mutants which have a higher or lower binding affinity for a Pro412-related protein relative to a wild type form of that Pro412-related target protein.
- 20 In an exemplary embodiment, a target protein can be provided as an immobilized protein (a "target"), such as by use of GST fusion proteins and glutathione treated microtitre plates. A gene (a "sample" gene) is amplified from cells of a patient sample, e.g., by PCR, ligated into an expression vector, and transformed into an appropriate host cell. The recombinantly produced Pro412-related protein is then contacted with the immobilized target protein, e.g., as a lysate or a semi-purified
- 25 preparation, the complex washed, and the amount of target protein /Pro412-related proteins complex determined and compared to a level of wild-type complex formed in a control. Detection can be by, for instance, an immunoassay using antibodies against the wild-type form of the Pro412-related protein, or by virtue of a label provided by cloning the sample Pro412-related gene into a vector which provides the protein as a fusion protein including a detectable tag. For example, a
- 30 myc epitope can be provided as part of a fusion protein with the sample Pro412-related gene. Such fusion proteins can, in addition to providing a detectable label, also permit purification of the sample Pro412-related protein from the lysate prior to application to the immobilized target. In yet

another embodiment of the subject screening assay, the two hybrid assay, described in the appended examples, can be used to detect mutations in either a Pro412-related gene or Pro412-related target gene which alter complex formation between those two proteins.

- In one embodiment, the biological sample contains protein molecules from the test subject.
- 5 Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject. In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting a Pro412-related protein, mRNA, or genomic DNA, such that the
 - 10 presence of a Pro412-related protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of a Pro412-related protein, mRNA or genomic DNA in the control sample with the presence of a Pro412-related protein, mRNA or genomic DNA in the test sample. The invention also encompasses kits for detecting the presence of Pro412-related protein, mRNA or genomic DNA in a biological sample. For example, the kit can comprise a labeled compound or
 - 15 agent capable of detecting a Pro412-related protein or mRNA or genomic DNA in a biological sample; means for determining the amount of a Pro412-related protein member in the sample; and means for comparing the amount of Pro412-related protein in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect Pro412-related protein or nucleic acid.
 - 20 In certain embodiments, detection involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202, the disclosures of which are incorporated herein by reference in their entireties), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegren et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *PNAS* 91:360-364, the disclosures of which are
 - 25 incorporated herein by reference in their entireties), the latter of which can be particularly useful for detecting point mutations in the Pro412-related protein-gene (see Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682, the disclosure of which is incorporated herein by reference in its entirety). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic
 - 30 acid sample with one or more primers which specifically hybridize to a Pro412-related gene under conditions such that hybridization and amplification of the Pro412-related protein-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with
 - 35 any of the techniques used for detecting mutations described herein.

Genotyping assays for diagnostics generally require the previous amplification of the DNA region carrying the biallelic marker of interest. However, ultrasensitive detection methods which

do not require amplification are also available. Methods well-known to those skilled in the art that can be used to detect biallelic polymorphisms include methods such as, conventional dot blot analyzes, single strand conformational polymorphism analysis (SSCP) described by Orita et al., PNAS 86 : 2766-2770 (1989), the disclosure of which is incorporated herein by reference in its entirety, denaturing gradient gel electrophoresis (DGGE), heteroduplex analysis, mismatch cleavage detection, and other conventional techniques as described in Sheffield et al. (1991), White et al. (1992), and Grompe et al. (1989 and 1993) (Sheffield, V.C. et al., Proc. Natl. Acad. Sci. U.S.A 49:699-706 (1991); White, M.B. et al., Genomics 12:301-306 (1992); Grompe, M. et al., Proc. Natl. Acad. Sci. U.S.A 86:5855-5892 (1989); and Grompe, M. Nature Genetics 5:111-117 (1993), the disclosures of which are incorporated herein by reference in their entireties).

Further methods are described as follows.

In one aspect, the nucleotide present at a polymorphic site can be determined by standard sequencing methods. In other examples, microsequencing methods are used, and the nucleotide at a polymorphic site in a target DNA is detected by a single nucleotide primer extension reaction as described by Chen and Kwok (1997) and, Chen and Kwok (Nucleic Acids Research 25:347-353 1997) and Chen et al. (Proc. Natl. Acad. Sci. USA 94/20 10756-10761,1997), the disclosures of which are incorporated herein by reference in their entireties). A preferred method of determining the identity of the nucleotide present at an allele involves nucleic acid hybridization. Any hybridization assay may be used including Southern hybridization, Northern hybridization, dot blot hybridization and solid-phase hybridization (see Sambrook et al., Molecular Cloning - A Laboratory Manual, Second Edition, Cold Spring Harbor Press, N.Y., 1989), the disclosure of which is incorporated herein by reference in its entirety). Hybridization conditions should be sufficiently stringent that there is a significant difference in hybridization intensity between alleles, and preferably an essentially binary response, whereby a probe hybridizes to only one of the alleles. Methods and chips of various formats for use in detecting biallelic polymorphisms can be produced on a customized basis by Affymetrix (GeneChip), Hyseq (HyChip and HyGnostics), and Protogene Laboratories and also described in as described in PCT application Nos. WO 92/10092 and WO 95/11995; US patent No. 5,424,186; and EP 785280, the disclosures of which are incorporated herein by reference.

30

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only, and are not intended to be limiting unless otherwise specified

Example 1: Purification and Characterisation of a Pro412-related protein

Human plasma was fractionated by a series of multiple chromatography columns. The fractions
 5 were analysed by mass spectrometry. Five tryptic fragments were found to belong to the protein
 sequence of Figure 1 by tandem mass spectrometry: FPGAVDGATYILVMVDPDAPSR,
 HWLVTDIK, IQQQELSAQAPSPPAHSGFHR, YQFFVYLQEGK and VISLLPK.

Example 2: Pro412-related proteins possess a PEBP-like functionnal domain

10 The Pro412-related proteins of the invention display in their C-term part a sequence
 predicted to be folded into an active site like the canonical ones found in the PEBP proteins. This
 structure has been described to form a ligand binding pocket at one end of the central beta sheet
 (Banfield *et al.*, *supra*). Interestingly, within this ligand-binding site, the Glu83 in the human PEBP
 protein (highlighted in Figure 2), which has been shown to adopt a relatively rare *Cis* peptide bond
 15 conformation, is mutated to Phe in the Pro412-related proteins of the invention, which is a residue
 not commonly found in a *Cis* conformation. In the mouse homolog of the Pro412-related proteins
 of the invention, this residue is mutated to Tyr (see Figure 3).

***Example 3: Pro412-related proteins possess a secretion signal peptide and a conserved,
 20 specific N-terminus domain***

The N-ter part of the Pro412-related proteins of the invention is found to correspond to a
 motif conserved across species, which comprises a very clear signal peptide for secretion of the
 protein and a conserved repetition of 4 Cys residues (see Figure 3). Computer modelisation of this
 conserved structure shows that it folds into a stable tertiary structure.

Example 4: Tissue distribution of Pro412-related proteins

25 Searches were conducted to establish the tissue distribution for two Pro412-related
 proteins of the invention, using the ESTs database dbEST (available at
<http://www.ncbi.nlm.nih.gov/dbEST/>). For the protein of SEQ ID NO:2, the distribution was as
 follows: out of 16 total ESTs, 7 were from pectoral muscle tissue (after mastectomy) (accession
 30 numbers: F20527, F28028, F33888, F36903, F25113, F26233, F26013); two were normal lung
 tissue (BF846558, BM984582) and three from carcinoma lung tissue (AA903899, AA902491,
 AI863957); and finally two were from healthy other tissues (H83927, AI218954) and two from
 tumorigenic other tissues (AA854779, BG214764).

35 For the protein of SEQ ID NO.4, the distribution was as follows: out of 8 total ESTs, 3
 were from normal prostate (AA688029, AA661735, AI659751), 2 were from carcinoma lung tissue
 (AI857902, AI816715), 1 was from skeletal muscle (AA192427) and 2 from other tissues
 (BE044451, AW975331).

Example 5: Chemical Synthesis of a Pro412-related protein

In this example, a Pro412-related protein of the invention is synthesized. Peptide fragment
 5 intermediates are first synthesized and then assembled into the desired polypeptide.

A Pro412-related protein can initially be prepared in, e.g. 5 fragments, selected to have a Cys residue at the N-terminus of the fragment to be coupled. Fragment 1 is initially coupled to fragment 2 to give a first product, then after preparative HPLC purification, the first product is coupled to fragment 3 to give a second product. After preparative HPLC purification, the second
 10 product is coupled to fragment 4 to give a third product. Finally, after preparative HPLC purification, the third product is coupled to fragment 5 to give the desired polypeptide, which is purified and refolded.

Thioester formation. Fragments 2, 3, 4, and 5 are synthesized on a thioester generating resin, as described above. For this purpose the following resin is prepared: S-acetylthioglycolic acid pentafluorophenylester is coupled to a Leu-PAM resin under conditions essentially as described by Hackeng et al (1999). In the first case, the resulting resin is used as a starting resin for peptide chain elongation on a 0.2 mmol scale after removal of the acetyl protecting group with a 30 min treatment with 10% mercaptoethanol, 10% piperidine in DMF. The N^a of the N-terminal Cys residues of fragments 2 through 5 are protected by coupling a Boc-thioproline (Boc-SPr, i.e.
 15 Boc-L-thioproline) to the terminus of the respective chains instead of a Cys having conventional N^a or S^b protection, e.g. Brik et al, J. Org. Chem., 65: 3829-3835 (2000).

Peptide synthesis. Solid-phase synthesis is performed on a custom-modified 433A peptide synthesizer from Applied Biosystems, using in situ neutralization/2-(1H-benzotriazol-1-yl)-1,1,1,3,3-tetramethyluronium hexafluoro-phosphate (HBTU) activation protocols for stepwise Boc chemistry chain elongation, as described by Schnolzer et al, Int. J. Peptide Protein Res., 40: 180-193 (1992). Each synthetic cycle consists of N^a-Boc -removal by a 1 to 2 min treatment with neat TFA, a 1-min DMF flow wash, a 10-min coupling time with 2.0 mmol of preactivated Boc-amino acid in the presence of excess DIEA and a second DMF flow wash. N^a-Boc-amino acids (2 mmol)
 20 are preactivated for 3min with 1.8mmol HBTU (0.5M in DMF) in the presence of excess DIEA (6mmol). After coupling of Gln residues, a dichloromethane flow wash is used before and after deprotection using TFA, to prevent possible high temperature (TFA/DMF)-catalyzed pyrrolidone carboxylic acid formation. Side-chain protected amino acids are Boc-Arg(p-toluensulfonyl)-OH, Boc-Asn(xanthyl)-OH, Boc-Asp(O-cyclohexyl)-OH, Boc-Cys(4-methylbenzyl)-OH, Boc-Glu(O-cyclohexyl)-OH, Boc-His(dinitrophenylbenzyl)-OH, Boc-Lys(2-Cl-Z)-OH, Boc-Ser(benzyl)-OH,
 25 Boc-Thr(benzyl)-OH, Boc-Trp(formyl)-OH and Boc-Tyr(2-Br-Z)-OH (Orpagen Pharma, Heidelberg, Germany). Other amino acids are used without side chain protection. C-terminal
 30 Fragment 1 is synthesized on Boc-Leu-O-CH₂-Pam resin (0.71 mmol/g of loaded resin), while for

Fragments 2 through 5 machine-assisted synthesis is started on the Boc-Xaa-S-CH₂-CO-Leu-Pam resin. This resin is obtained by the coupling of S-acetylthioglycolic acid pentafluorophenylester to a Leu-PAM resin under standard conditions. The resulting resin is used as a starting resin for peptide chain elongation on a 0.2 mmol scale after removal of the acetyl protecting group with a 5 30min treatment with 10% mercaptoethanol, 10% piperidine in DMF.

After chain assembly is completed, the peptide fragments are deprotected and cleaved from the resin by treatment with anhydrous hydrogen fluoride for 1hr at 0°C with 5% p-cresol as a scavenger. In all cases except Fragment 1, the imidazole side chain 2,4-dinitrophenyl (DNP) protecting groups remain on His residues because the DNP-removal procedure is incompatible 10 with C-terminal thioester groups. However DNP is gradually removed by thiols during the ligation reaction, yielding unprotected His. After cleavage, peptide fragments are precipitated with ice-cold diethylether, dissolved in aqueous acetonitrile and lyophilized. The peptide fragments are purified by RP-HPLC with a C18 column from Waters by using linear gradients of buffer B (acetonitile/0.1% trifluoroacetic acid) in buffer A (H₂O/0.1% trifluoroacetic acid) and UV 15 detection at 214nm. Samples are analyzed by electrospray mass spectrometry (ESMS) using an Esquire instrument (Brücker, Bremen , Germany), or like instrument.

Native chemical ligations. As described more fully below, the ligation of unprotected fragments is performed as follows: the dry peptides are dissolved in equimolar amounts in 6M guanidine hydrochloride (GuHCl), 0.2M phosphate, pH 7.5 in order to get a final peptide 20 concentration of 1-8 mM at a pH around 7, and 1% benzylmercaptan, 1% thiophenol is added. Usually, the reaction is carried out overnight and is monitored by HPLC and electrospray mass spectrometry. The ligation product is subsequently treated to remove protecting groups still present. The formyl group of Trp is cleaved by shifting the pH of the solution up to 9.0 with hydrazine and incubating for 1h at 37° C. Opening of the N-terminal thiazolidine ring further 25 required the addition of solid methoxamine to a 0.5M final concentration at pH3.5 and a further incubation for 2h at 37°C. A 10-fold excess of Tris(2-carboxyethyl)phosphine is added before preparative HPLC purification. Fractions containing the polypeptide chain are identified by ESMS, pooled and lyophilized.

The ligation of fragments 4 and 5 is performed at pH7.0 in 6 M GuHCl. The concentration 30 of each reactant is 8mM, and 1% benzylmercaptan and 1% thiophenol were added to create a reducing environment and to facilitate the ligation reaction. An almost quantitative ligation reaction is observed after overnight stirring at 37°C. At this point in the reaction, O-NH₂.HCl is added as a powder to a 0.1 M final concentration and hydrazine is added to shift the pH to 9.0, for the removal of the formyl group of any Trp residues. After a 1h incubation at 37°C, O-NH₂ HCl is 35 further added to the solution to get a 0.5M final concentration, and the pH adjusted to 3.5 in order to open the N-terminal thiazolidine ring. After 2h incubation at 37°C, ESMS is used to confirm the completion of the reaction. The reaction mixture is subsequently treated with a 10-fold excess of

Tris(2-carboxyethylphosphine) over the peptide fragment and after 15min, the ligation product is purified using the preparative HPLC (e.g., C4, 20-60% CH₃CN, 0.5% per min), lyophilised, and stored at -20°C.

The same procedure is repeated for the remaining ligations with slight modification.

- 5 Whenever the ligation takes place at an Ile-Cys or Val-Cys site, the ligation reaction is extended to 48h.

Polypeptide Folding. The full length peptide is refolded by air oxidation by dissolving the reduced lyophilized protein (about 0.1 mg/mL) in 1M GuHCl, 100mM Tris, 10mM methionine, pH 8.6 After gentle stirring overnight, the protein solution is purified by RP-HPLC as described
10 above.

Example 6: Preparation of antibody compositions

Substantially pure Pro412-related protein or a portion thereto is obtained. The
15 concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms per ml. Monoclonal or polyclonal antibodies to the protein can then be prepared as follows: Monoclonal Antibody Production by Hybridoma Fusion Monoclonal antibody to epitopes in the Pro412-related protein or a portion thereof can be prepared from murine hybridomas according to the classical method of Kohler and
20 Milstein (Nature, 256: 495, 1975) or derivative methods thereof (see Harlow and Lane, *Antibodies A Laboratory Manual*, Cold Spring Harbor Laboratory, pp. 53-242, 1988), the disclosure of which is incorporated herein by reference in its entirety.

Briefly, a mouse is repetitively inoculated with a few micrograms of the Pro412-related protein or a portion thereof over a period of a few weeks. The mouse is then sacrificed, and the
25 antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the
30 supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, E., Meth. Enzymol. 70: 419 (1980), the disclosure of which is incorporated herein by reference in its entirety. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis, L. et al. *Basic Methods in Molecular Biology* Elsevier, New York. Section 21-2, the
35 disclosure of which is incorporated herein by reference in its entirety.

Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes in the Pro412-related protein or a portion thereof can be prepared by immunizing suitable non-human animal with the Pro412-related protein or a portion thereof, which can be unmodified or modified to enhance immunogenicity. A suitable nonhuman animal, preferably a non-human mammal, is selected. For example, the animal may be a mouse, rat, rabbit, goat, or horse. Alternatively, a crude protein preparation which has been enriched for Pro412-related proteins or a portion thereof can be used to generate antibodies. Such proteins, fragments or preparations are introduced into the non-human mammal in the presence of an appropriate adjuvant (e. g. aluminum hydroxide, RIBI, etc.) which is known in the art. In addition the protein, fragment or preparation can be pretreated with an agent which will increase antigenicity, such agents are known in the art and include, for example, methylated bovine serum albumin (mBSA), bovine serum albumin (BSA), Hepatitis B surface antigen, and keyhole limpet hemocyanin (KLH). Serum from the immunized animal is collected, treated and tested according to known procedures. If the serum contains polyclonal antibodies to undesired epitopes, the polyclonal antibodies can be purified by immunoaffinity chromatography.

Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. Techniques for producing and processing polyclonal antisera are known in the art, see for example, Mayer and Walker (1987), the disclosure of which is incorporated herein by reference in its entirety. An effective immunization protocol for rabbits can be found in Vaitukaitis, J. et al. *J. Clin. Endocrinol. Metab.* 33: 988-991 (1971), the disclosure of which is incorporated herein by reference in its entirety. Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. et al., Chap. 19 in: *Handbook of Experimental Immunology* D. Wier (ed) Blackwell (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12: M). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: *Manual of Clinical Immunology*, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D. C. (1980).

Antibody preparations prepared according to either the monoclonal or the polyclonal protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; or they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. The antibodies may also be used in

60394576 .07080P

therapeutic compositions for killing cells expressing the protein or reducing the levels of the protein in the body.

CLAIMS

1. A substantially purified protein having an amino acid sequence of SEQ ID NOS 1 to 4; and
5 biologically active portions thereof.
2. A substantially purified protein consisting of an amino acid sequence of SEQ ID NOS 1 to 4;
and biologically active portions thereof.
- 10 3. A Pro412-related protein according to claim 1, wherein said protein comprises at least one
amino acid deletion, substitution or insertion with respect to said amino acid sequence of claims 1
or 2.
- 15 4. A substantially purified protein comprising a Pro412-related protein according to claims 1 or 2,
wherein said Pro412-related protein is fused to another, heterologous protein or amino acid
sequence.
- 20 5. A method of making a Pro412-related protein, said method comprising
providing a population of host cells comprising a recombinant nucleic acid encoding a
Pro412-related protein of Claims 1 or 4; and
culturing said population of host cells under conditions conducive to the expression of said
recombinant nucleic acid;
whereby said protein is produced within said population of host cells.
- 25 6. The method of claim 5, further comprising purifying said protein from said population of cells.
7. An antibody that selectively binds to the protein of Claims 1 to 3.
- 30 8. An antibody according to Claim 7, wherein said antibody is capable of inhibiting binding of said
protein to a Pro412-related substrate.
- 35 9. A method of determining whether a Pro412-related protein is expressed within a biological
sample, said method comprising the steps of:
 - (a) contacting a biological sample from a subject with a detectable protein that selectively
binds to the protein of Claims 1 or 2; and
 - (b) detecting the presence or absence of binding of said detectable protein to a protein within
said sample;wherein a detection of said binding indicates that said Pro412-related protein is expressed

within said sample.

10. The method of claim 11, wherein said detectable protein is an antibody.

5 11. A method of determining whether a mammal has an elevated or reduced level of Pro412-related protein expression, said method comprising the steps of:

(a) providing a biological sample from said mammal; and

(b) comparing the amount of a Pro412-related protein of SEQ ID NOS 1 to 4 within said biological sample with a level detected in or expected from a control sample;

10 wherein an increased amount of said Pro412-related protein within said biological sample compared to said level detected in or expected from said control sample indicates that said mammal has an elevated level of Pro412-related protein expression, and wherein a decreased amount of said Pro412-related protein within said biological sample compared to said level detected in said control sample indicates that said mammal has a reduced level of Pro412-related protein expression.

15 12. A method of identifying a candidate Pro412-related protein modulator, said method comprising:

(a) contacting a Pro412-related protein according to Claims 1 or 2 with a test compound; and

20 (b) determining whether said compound selectively binds to said protein;

wherein a determination that said compound selectively binds to said protein indicates that said compound is a candidate inhibitor of a Pro412-related protein, a candidate Pro412-related protein modulator.

25 13. A method of identifying a candidate Pro412-related protein modulator, said method comprising:

(a) contacting said Pro412-related protein of Claims 1 or 2 with a test compound; and

(b) determining whether said compound selectively modulates interaction with a Pro412-related protein substrate protein,

30 wherein a determination that said compound selectively modulates said interaction indicates that said compound is a candidate Pro412-related protein modulator.

14. A method of identifying a candidate Pro412-related protein modulator, said method comprising:

35 (a) contacting said Pro412-related protein of Claims 1 or 2 with a test compound; and

(b) determining whether said compound selectively inhibits a biological activity of said Pro412-related protein,

60394576 ,070802

wherein a determination that said compound selectively inhibits a Pro412-related activity indicates that said compound is a candidate Pro412-related protein inhibitor.

60394576 . 070302

Figure 1

	<u>MGWTMRLVTAALLGLMMVVTGDEDENSPCAHEALLDETLFCQGLEVFYPELGNIGCKV</u>	60
	<u>VPDCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHVLVTDIKG</u>	120
5	<u>ADLKEGKIQQQELSAQAPSPPAHSGFHRYQFFVYLQEGKVISLLPKENKTRGSWKMDRF</u>	180
	<u>LNRFHGEPEASTQFMTQNYQDSPTLQAPRERASEPKHKNOAEIAAC</u>	227

68394576 .070802

Figure 2

	pdblbeh pdbl44 Q96S96	VDSLKWGPISLQEVD EQPQHPLHV TYAGAAVDELGKVLTPTQVKNRPTS VDSLKWGPISLQEVD ERPQHPLQVKYCGAEVDELGKVLTPTQVKNRPTS EDENSPCAHEALLD EDTLFCQGLEVFP E LGNIGCKVVPDCNNYRQKITS
5	pdblbeh pdbl44 Q96S96	< beta > < beta > ----ISWDGLDSGKLYTLVLTDPDAPS R KDPKYREWHHFLVVNMKGNDI ----ITWDGLDPGKLYTLVLTDPDAPS R KDPKYREWHHFLVVNMKGNNI WMEPIVKFPGAVDGAT YILVMVDPDAPSRAEPRQRFWRH WLVT DIKGADL
10	pdblbeh pdbl44 Q96S96	<beta> < beta > S----SGTVLSDYVGSGPPKG TGLHRYVWL VYEQDRPLKCDEPILSNRSG S----SGTVLSDYVGSGPPKG TGLHRYVWL VYEQEGPLKCDEPILSNRSG KECKTQGQELSAYQAPSPPAHSGFHRYQFFVYLQEGKVISLLPKEN----
15	pdblbeh pdbl44 Q96S96	< alpha > < beta > DHRGKFVKVASFRKKYELRAPVAGTCYQAEWDDYVPKLYEQLSG DHRGKFVKVASFRKKYELGAPVAGTCYQAEWDDYVPKLYEQLSG KTRGSWKMDRFLNRFHLGEPEASTQFMTQNYQDSPTLQAPRER
20		

160394576 070802

Figure 3

5	Q96S96 Q8WW74 Q9D9G2 Q9D9L9	DEDENSPCAHEALLDEDTLFCQGLEVFYPELCNIGCKVVPDCNNYRQKIT DEDENSPCAHFAALLDEDTLFCQGLEVFYPELGNIGCKVVPDCNNYRQKITCFLPLPKEDVSI.CRNLEVLFYMEMGNISCKIVPKCNLYRQKITPCFLPLPKEDVSLGRNLEVLFYMEMGNISCKIVPKCNLYRQKIT
10	Q96S96 Q8WW74 Q9D9G2 Q9D9L9	SWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHVLVTDIKGAD SWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHVLVTDIKGAD AWQAPIVKFHTALDGALYLLVMVDPDAPSRSNPVMKYWRHWLVSNTGAD AWQAPIVKFHTALD.....
15	Q96S96 Q8WW74 Q9D9G2 Q9D9L9	LKEGKIQGQELSAYQAPSPPAHSGFHRYOFFVYLQEGKVISLLPKENKTR LKEGKIQGQELSAYQAPSPPAHSGFHRYOFFVYLQEGKVISLLPKENKTR MKSGSIRGNVLSDYSPPTPPPETGVHRYQFFVYLQGDKDLSVEEKANL
20	Q96S96 Q8WW74 Q9D9G2 Q9D9L9	GSWKMDRFLNRFHLGEPEASTQFMTQNYQDSPTLQAPRERASEPKHKNQA GSWKMDRFLNRFHLGEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTR. GGWNLDKFLQQYGLRDPDTSTQFMTQ.....
25	Q96S96 Q8WW74 Q9D9G2 Q9D9L9
30	Q96S96 Q8WW74 Q9D9G2 Q9D9L9	EIAAC

35

60394576 .070802

SEQUENCE LISTING

<110> Geneprot, Inc.
5 <120> secreted proteins
<130>
10 <160> 18
<170> PatentIn version 3.0
15 <210> 1
<211> 227
<212> PRT
<213> Homo Sapiens
20 <220>
<221> misc_feature
<222> (1)..(227)
<223> entire protein, 'erase' form
25 <220>
<221> SIGNAL,
<222> (1)..(22)
<223> predicted by SignalP version 2.0
30 <220>
<221> SITE
<222> (211)..(211)
<223> predicted variation site
35 <220>
<221> misc_feature
<222> (220)..(227)
40 <223> sequence related to the variation site 211
45 <220>
<221> DOMAIN
<222> (90)..(112)
<223> matching PROSITE pattern PBP PS01220
(phosphatidylethanolamine-binding protein)
50 <400> 1
Met Gly Trp Thr Met Arg Leu Val Thr Ala Ala Leu Leu Leu Gly Leu
} 5 10 15
Met Met Val Val Thr Gly Asp Glu Asp Glu Asn Ser Pro Cys Ala His
20 25 30
55 Glu Ala Leu Ile Asp Glu Asp Thr Leu Phe Cys Gln Gly Leu Glu Val
35 40 45
60 Phe Tyr Pro Glu Leu Gly Asn Ile Gly Cys Lys Val Val Pro Asp Cys
50 55 60
Asn Asn Tyr Arg Cln Lys Ile Thr Ser Trp Met Glu Pro Ile Val Lys
65 70 75 80
65 Phe Pro Gly Ala Val Asp Gly Ala Thr Tyr Ile Leu Val Met Val Asp
85 90 95
Pro Asp Ala Pro Ser Arg Ala Glu Pro Arg Gln Arg Phe Trp Arg His
100 105 110

Trp Leu Val Thr Asp Ile Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile
 115 120 125
 5 Gln Gly Gln Glu Leu Ser Ala Tyr Gln Ala Pro Ser Pro Pro Ala His
 130 135 140
 Ser Gly Phe His Arg Tyr Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys
 145 150 155 160
 10 Val Ile Ser Leu Leu Pro Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys
 165 170 175
 Met Asp Arg Phe Leu Asn Arg Phe His Leu Gly Glu Pro Glu Ala Ser
 180 185 190
 Thr Gln Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala
 195 200 205
 20 Pro Arg Glu Arg Ala Ser Glu Pro Lys His Lys Asn Gln Ala Glu Ile
 210 215 220
 Ala Ala Cys
 225
 25 <210> 2
 <211> 205
 <212> PRT
 <213> Homo Sapiens
 30 <220>
 <221> misc_feature
 <222> (1)..(205)
 <223> mature peptide
 35 <400> 2
 Asp Glu Asp Glu Asn Ser Pro Cys Ala His Glu Ala Leu Leu Asp Glu
 1 5 10 15
 40 Asp Thr Leu Phe Cys Gln Gly Leu Glu Val Phe Tyr Pro Glu Leu Gly
 20 25 30
 Asn Ile Gly Cys Lys Val Val Pro Asp Cys Asn Asn Tyr Arg Gln Lys
 45 35 40 45
 Ile Thr Ser Trp Met Glu Pro Ile Val Lys Phe Pro Gly Ala Val Asp
 50 55 60
 50 Gly Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp Ala Pro Ser Arg
 65 70 75 80
 Ala Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu Val Thr Asp Ile
 85 90 95
 55 Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile Gln Gly Gln Glu Leu Ser
 100 105 110
 60 Ala Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe His Arg Tyr
 115 120 125
 Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser Leu Leu Pro
 130 135 140
 65 Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys Met Asp Arg Phe Leu Asn
 145 150 155 160
 Arg Phe His Leu Gly Glu Pro Glu Ala Ser Thr Gln Phe Met Thr Gln
 165 170 175

60394576 070802

Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala Pro Arg Glu Arg Ala Ser
180 185 190

5 Glu Pro Lys His Lys Asn Gln Ala Glu Ile Ala Ala Cys
195 200 205

10 <210> 3
<211> 223
<212> PRT
<213> Homo Sapiens

15 <220>
<221> misc_feature
<222> (1)..(223)
<223> entire protein, 'grase' form

20 <220>
<221> STGNAL
<222> (1)..(22)
<223> predicted by SignalP version 2.0

25 <220>
<221> SITE
<222> (211)..(211)
<223> predicted variation site

30 <220>
<221> misc_feature
<222> (220)..(223)
<223> sequence related to the variation site 211

35 <220>
<221> DOMAIN
<222> (90)..(112)
40 <223> matching PROSITE pattern PBP PS01220
(phosphatidylethanolamine-binding protein)

45 <400> 3

45 Met Gly Trp Thr Met Arg Leu Val Thr Ala Ala Leu Leu Gly Leu
1 5 10 15

50 Met Met Val Val Thr Gly Asp Glu Asp Glu Asn Ser Pro Cys Ala His
20 25 30

55 Glu Ala Leu Leu Asp Glu Asp Thr Leu Phe Cys Gln Gly Leu Glu Val
35 40 45

55 Phe Tyr Pro Glu Leu Gly Asn Ile Gly Cys Lys Val Val Pro Asp Cys
50 55 60

60 Asn Asn Tyr Arg Gln Lys Ile Thr Ser Trp Met Glu Pro Ile Val Lys
65 70 75 80

65 Phe Pro Gly Ala Val Asp Gly Ala Thr Tyr Ile Leu Val Met Val Asp
85 90 95

65 Pro Asp Ala Pro Ser Arg Ala Glu Pro Arg Gln Arg Phe Trp Arg His
100 105 110

Trp Leu Val Thr Asp Ile Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile
115 120 125

60391576 . 070802

Gln Gly Gln Glu Leu Ser Ala Tyr Gln Ala Pro Ser Pro Pro Ala His
130 135 140

Ser Gly Phe His Arg Tyr Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys
5 145 150 155 160

Val Ile Ser Leu Leu Pro Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys
165 170 175

10 Met Asp Arg Phe Leu Asn Arg Phe His Leu Gly Glu Pro Glu Ala Ser
180 185 190

Thr Gln Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala
195 200 205

15 Pro Arg Gly Arg Ala Ser Glu Pro Lys His Lys Thr Arg Arg Arg
210 215 220

<210> 4
20 <211> 201
<212> PRT
<213> Homo Sapiens

<220>
25 <221> misc_feature
<222> (1)..(201)
<223> mature peptide

30 <400> 4

Asp Glu Asp Glu Asn Ser Pro Cys Ala His Glu Ala Leu Leu Asp Glu
1 5 10 15

35 Asp Thr Leu Phe Cys Gln Gly Leu Glu Val Phe Tyr Pro Glu Leu Gly
20 25 30

Asn Ile Gly Cys Lys Val Val Pro Asp Cys Asn Asn Tyr Arg Gln Lys
35 40 45

40 Ile Thr Ser Trp Met Glu Pro Ile Val Lys Phe Pro Gly Ala Val Asp
50 55 60

Gly Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp Ala Pro Ser Arg
45 65 70 75 80

Ala Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu Val Thr Asp Ile
85 90 95

50 Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile Gln Gly Gln Glu Leu Ser
100 105 110

Ala Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe His Arg Tyr
115 120 125

55 Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser Leu Leu Pro
130 135 140

Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys Met Asp Arg Phe Leu Asn
60 145 150 155 160

Arg Phe His Leu Gly Glu Pro Glu Ala Ser Thr Gln Phe Met Thr Gln
165 170 175

65 Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala Pro Arg Gly Arg Ala Ser
180 185 190

Glu Pro Lys His Lys Thr Arg Arg Arg
195 200

```

<210> 5
<211> 12042
<212> DNA
5 <213> Homo Sapiens

<220>
<221> misc_feature
<222> (1)..(12042)
10 <223> public sequence NT_008300.6 (545299..557340 c)

<220>
<221> exon1
<222> (1952)..(2082)
15 <223> coding exon

<220>
<221> CDS
<222> (1952)..(2080)
20

<220>
<221> exon2
<222> (9356)..(9482)

25 <220>
<221> CDS
<222> (9357)..(9482)

<400> 5
30 gtttcaaatt ccaggtctgc cacttcttgg tgaccaagat caggtttaa ctaatgtcac 60
      tggcgtgtct ctcctccat aagatgggt gataataqtg cctattcca aggattggca 120
      tgaggattcg gagataaaag tgccctggcac gtggagtaag gaagaaccgc cactggaagt 180
35 gggctcaaca gcaqcaacct gcactgtggg ctctcccttc ttccagggcc ccatcaactgt 240
      gctaaaggag ttgacaccca gaaaagttaa gacttctcta gtggatgcc 111ggaaat 300
40 gagtttctllg gtcgtggca gcccttctgg cagcatctcc ttggctgtgg ctgtcctgtg 360
      agctgcatgc tcaacctata ccccagggcc caacatgggg caggcacgtg ggaggtgctt 420
      aatgcctatg agtgaagtgg gttgcaaagg aatgaatgaa tggatggatc cgtgagagaa 480
45 tacttagact attcataaat tattgagcat ttactatgtg ccaggcacct tactgaggc 540
      ttaacctggg tgatttcatt ctcataatga ttttgtgagg aggctactat tagactctgt 600
50 ccactagggg agatcttggg aatatgtttt ctggttgggg caagggggag tcctggttgt 660
      acttcagaga acctaataqat tgagaaatgt gtttaataatg tagccccctaa ggggggcttt 720
      ttgctacagg ttgggatctt ggaggcagga attgagccag aaagcagact atgctgaaag 780
55 tagtttctt aaaggccttgc ctgcgtgtac ttacagcga tgatctcata ttccttagca 840
      aggcaccacg gtgacaggaa ttatttact tggatgtgaa aqgaaaaagt ggtgcattct 900
60 aaggagacac aglggcctcc cagcatccag agagagctca gacgacaagc caggagtcct 960
      gagttgctgg cacqtccttc craccccccgc ctctgagaag cagagagcag agctattgga 1020
      aagcctgttgc aagccgcct ttattcttctt ccctgccac tcaggctcaa accagagtcc 1080
65 cacatgcacc tgqlggcagg actgttcccc aaggaaalgc aggccacagt ccccggtcc 1140
      ctttccccc agtgcaclgc tggggcaggc cacccttcc tgagccccaa gcaagtcctcc 1200

```

60394576 070802

	agacccctt ctccagatac gattatggg ggagagtgt gagtgcagca cagctgcagg	1260
	gtgcacccag acgcacccctt ctgtcctcag ctggcccttc ggagctggcc cttcggagct	1320
5	ggcctcggtc tgggggtgcac ctaccccccac cccctgcagc agtccctctt acacaaggta	1380
	actgtgtcca accaccccca ccttggggaa gtcacaggc tgtggagaca ggtgggtggaa	1440
10	gaatgggggc caggagagtc agagccggtc tgaggggcac agcccaactc ctgcctgaa	1500
	gcctggggtc accagcttcc cagggtttgc cttacagccca aggagcaggc tggcctttgg	1560
	ccagctgacc atccctggc caacagcagg gactggacaa ggagctgcac cgaggagctt	1620
15	atataagcac ttctactttt cagccctqctc tctgccttg cagccaggac aagagtgaca	1680
	acaaagctgt gtccctgccag gggagggggcc tctgccagga acgtacattt agtccaaactt	1740
20	ggaagaggcc cagtcctccc aaagtacttg tgcctgggtt gtggacttggaa ttatctgcgg	1800
	agccctggaa gtcctgttc ctctccctt tgcttaacca gaggtctcta tcctgcctaa	1860
	tggggccctt gggtcggggg cacaggtgtt gggcaggaag gccctgcag gggatccagc	1920
25	cgccttaaag ttcttcgtt tccaggtgcc c atg ggt tgg aca atg agg ctg Met Gly Trp Thr Met Arg Leu	1972
	1 5	
30	gtc aca gca gca ctg tta ctg ggt ctc atg atg gtc act gga gac Val Thr Ala Ala Leu Leu Leu Gly Leu Met Met Val Val Thr Gly Asp	2020
	10 15 20	
35	gag gat gag aac agc ccg tgc cat gag gcc ctc ttg gac gag gac Glu Asp Glu Asn Ser Pro Cys Ala His Glu Ala Leu Asp Glu Asp	2068
	25 30 35	
40	acc ctc ttt tgc cagtaagtca gggcaggctt ggaggcggqa qqcacaaagg Thr Leu Phe Cys	2120
45	ctaccaggca ggttagatggc agggagggt atatccagaa ccaagaagca ccatctcttg agcctgatt tatggaaagg cagaacacgt gtcagatctg gagaccacct tgatggagga	2180
	2240	
50	tggcaactgag aaagggattt agaaaacagac tcccttcgtc cttccagaat gcacctctcc caggaccagg agggctgtgt atgtggtcgg ctctgtttgg gcccctccaa cctgggggg	2300
	2360	
55	ctttgtgggg gcaggaggct gagaccctca gagagagacc cagagagaga ccatccctcc ctgcacccctg aatgccttcc cccccccagaa ccccaacaaa gaggcatttg aatgaaagaa	2420
	2480	
60	ggtgttatgt tagacccttc ctgtgtgaga qgttttqtc aqgtatagga cccaggagcg tttgcagg gcccctttagg gatcatggca ctggatggag ttgggtgacg aggctggctt	2540
	2600	
	cttcccttcc aaggccacaa gctgcttcta ctttggtaa gctgaaagc ctggcggtcg	2660
	gctctctctt caccactct gccggttaga aaagcaagcc ttcatttgtg tggacttgg ggtcacccctg aggaclcccc aggcacaaact gatgcacaaac ctccclgcca gtttgggtggc	2780
	2840	
65	tgaggagggg gacacccac tggcccacg tgtqaaaaal qaqaqaggag gaaaaagaac tttgactact tactctcttc tgtacctttt aatttcaca aacattactg ttcatgaaaa	2900
	ttttgttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2960
	ataaaatacat gtcctctttt aaaattatgt agttgtatcc ttctaaaaaa gtaacatttt	3020

	tgtlgggtga agtggctcac atgtgtatc ccagcacttt gggaggccaa ggtggaggat	3080
5	cacttgagcc caggggttt agaccaccc gggcaacata gagagactct tttctctaca	3140
	aaacagttt aaaaacttagcc aagcatgtg atgcacaccc gtcgtcccag ctactcaaga	3200
	ggttgaggttgg ggaggatcgc tttaaaaaaag yaggtcaagg ctggctctgag ccatgattgc	3260
10	gccactgcat tccagcctgg gcaacagagc cagaccctgc ctaaaacaaca acaacaacaa	3320
	caacaaaaaag taacaaactt tgttcggttc agtatatttt gaaaaatctt gttaaatagt	3380
15	tctgggtgtt ggctggcgc ggtggctcac acctgtatc ccaagacttt gggagactga	3440
	ggcaggcaga taacaaggc aggagtttga gaccagctg accaacacgg tgaaaaccctg	3500
	tctctaclaa aaataaaaaa aaatttagccg ggcattggtgg cgcacacctg taatcccagc	3560
20	tactcaggag gctgaggcag gagaatccc tqaatctggg aggtggaggt tgcagtgagc	3620
	cgagattgcg ccactgcact ctgcctggg cgacagacca agactctgtc taaaaaaaaaca	3680
25	aacaaaaaaaaa caancaagca agaaaaacagt lctgggtgaa gtttacacgt ggaaggagga	3740
	agagaaaaacgt ggttgcctct cacaagtggg atctaatttt ctatgtgtc tgctgagtga	3800
	gclgaggggc tggcaggctt ctgcgtctga gcagtgcag gtcaacacccc ccgtaaaccag	3860
30	gagaggtcag ctggggcagg gtacaqcagg cctggggcac aataagggcc ccttaaggag	3920
	gacctcagca gttcctggg tttccacac tgtggaaat gaggatgtct gtagggtcat	3980
35	tgggggtgac cagccclgagg ctccagatgc cctaggctct gccctgccaa gatcacatgt	4040
	tggtgccctg gtgaccatc catggggaca cacgtggaa gtcccaggc accttctgtg	4100
	tgccaggtgc tctgccccgtg cctggactcc ttgactcctt tcactgtcag taacgaacct	4160
40	acaaaatctg tgtagcacc cccattctgc agatgaagaa tccccagaga ttaagaaatc	4220
	tgcccacagt gtttacatag cttaaaaagta gcagagccaa gatctagccc ctttgtctga	4280
45	ctccacactc acgtcttttc cctatctgtc tctgacaacc ttccccccag gttcagggtct	4340
	cccgactcg agtcaagggg gccccatctc aggactccca gaatacacca gcaggcccc	4400
	agtctgagaa agagacttat gacctaagg ttgcattgcca ggaagatgtc ttttgatttc	4460
50	aagccatcat cgcaaaaagay gagctttgtc tttattgtca ctgaagattt ttgtttggaa	4520
	gagagcatca tgaatgatga tgaaatgaaa aatqcatcaa ggacagatcg gacccctgag	4580
55	gagacagcag gacccctctgc actgaccctt cctcaagatc cccactcgcc tcaggctctg	4640
	accttggccaa gagatcccc agagaaggca ctccclgtt gggcatttgg tctaccaggg	4700
	aactcttaaa gattggactc aatgccttaaa ggtattaatc tataagggag qaattttal	4760
60	tactgtgaca agctgtgaga cattgttaag agtaacgtt ctcggaaatc agaccttccc	4820
	tgcttctgag ataccccaga caagaccagc aaaaaccaccc acggtcccaa ccctgggtc	4880
65	agtcgcagaa tggaaagaac catgattctc caaacatgcg gccaggaggt gaatccactc	4940
	aaaaaaccaccc acctgtgtct gccalccagt caqccclgca agctccctta gggatgggaa	5000
	ctgctgtca aagcgtccctg ctttttgggtg ttgaaaagct ccgttcttca gaaagtacat	5060

	tcttaccccg aggccctgct gctcltcgg aacaccaacc ctacccgttg tagtgccggat	5120
	ccacgagggt cacaaactag tctatgtttc catcttctt gtggcatttc atgtgtgtgg	5180
5	tgacagggtgc catgcctgc cctgcctt catccccagc tcccccaac ctggcagtc	5240
	atccccagtc cttcaaccctc tctcaggggc cttggctctt cctgttaaaa tacatttctt	5300
10	ttaqagacag agtcttgtct gtcacttagg ctggagtgcg gtggcctgat catagtcac	5360
	tgcagcctt aacacctggg ctcaagcaat cctctgcct caacgtcctg agtagctggg	5420
	actacagggtg cacgccacca caccctggcta agtttcttat tttttgtaga gttgggagtc	5480
15	tcggatqta gtccaggctg ttctcaaact cctgggctca agcaatcctg ctgcctcage	5540
	ctccccaaagt gccggattta caggcatgca ctgcataacc tggcttgctg ttttcaagt	5600
20	gggggtgaccc acactttcac agtgagcccc ctcttcgagg cagcctcata tcccattgtgt	5660
	ctttttgtca tcctcactgt cttatgtctg tcaagctacca taacaaaata ccttagacta	5720
	ggtcatttctt aaacagcagg aatgtattac tcacaactct ggaggctgag aagcgccaaa	5780
25	atcaagggtgc tagtggattc agtgtgcggc acaagggtgc tctctgcctt aaagatggca	5840
	ccttcttgct gcgacttcac atgggggaag ggaagggcag ctcccttcag ccclclcltat	5900
30	gagggtgcct tttataaggg tgacttaatc acctcttcca agtcccgcctt gtaataacca	5960
	tcatatttag gattaaatcc caacatataa attttgagga gacacaaaca cttagaccat	6020
	agcactcact gaggggcagc tgaaattcta agcccttcat acaggcacag ctgagaagcc	6080
35	acgtttctag agacagggtc tcagtgcate accctctgac acccaggctg gaatgtagaa	6140
	gcaggatcat ggctcaactgc aaccttgacc tcccaqactc aagcgatcct cccacctcag	6200
40	cctcccatat agctggacta cagatgcgtt ccaccacact tggctaattt ttctatgttt	6260
	ttatagagac aggatctcar taatgtgccc aggctggct ctaactcctg ggctcaagt	6320
	atcatctcac tcagcctcc caaagtgcgtt ggattacagg cttagccac cacgccccagc	6380
45	ccttgcatac laaltcttgc catgacagag acccacatga ccacccatgtt caaactgaag	6440
	cctggaggaa tgaaacagct cactaaggaa aacgtgtggt gggggaaaga aagtgcagcc	6500
	aagccccacag tactgttct ctgttgtaca caacagatca ctccaaacacg tagcagtc	6560
50	aaataggtat tgcataacag tttctcagtt actcatctgg agataggaat agtgataaca	6620
	gctgcctcaa ggtatgggtt catgaaatga aatcatacat atgtgtatgc cccgatatat	6680
55	lggggtggcc cagtaaaagt catgccttc atccccagtc cttatgttag cattagttca	6740
	tggccaaagct gggccaggc cccctqatct tttctcataca ctgtctagag tgccctcaag	6800
60	tcttgcataag gagatcagga gatccccaaa ctctccacat ggaaatttaga agaattccat	6860
	gtcaggtaaa ccacaaggcag caagggcaca caggaagggt gcacccttga ggttttctgg	6920
	cttgaatgca agttgtttt gagatagggc ttqctctgt cacccagggt ggagtgcgggt	6980
65	gatgcaatca tggctcaactg cagcttctac ctccctggct caagtgcattcc tcccaagtag	7040
	ctgggaacac aggccaggcac tarcatgggt tttttttt glllgttlql ltttttaatt	7100
	tttggtagaa acagggtctc actatgtgc cttaggtggc ctcagattcc tgggctcaag	7160

603945アセ・ロアロロ正

	lgalgclccc accccagcct cccaaagagc ttagattata ggcatgagcc accgtgccta	7220
5	qtcttqaatg caaaaatttaa agtgtatttt ctctactttaatttttg attcagctca	7280
	cctcagggtt gtccagggtcg tctcgtaagt caatttgatc atcagaatttgggttacca	7340
10	gclctgtgcc atctgcagat caaataaaattt tgtctccgtg tccttctcca agactttgttgc	7400
	aaagcagctg agcagggcag agqcctgcac atgccttccc cccatcatcc taggtggggc	7460
	aatacgttat gttcaaaccgc tggcctggca gtgcctatg ggcgtgaaga gagggacaga	7520
15	gacaagaccc aaatccctgg ttccctggca tgtgcctgttgc cttcagggc actctctgtg	7580
	gctctgtgcc cacctcatca aataacaattt tgatgatttt gcttgtttaa tcctttgtat	7640
	attacaaca attccactt gataaaaagg ggagagggtt tatgggcgtg agggcgtcccc	7700
20	tccaccctca gtttacctcc tccctacctc tcagtttcc catcctgtgc tgtttgcgtgt	7760
	ctgtctctct gcccctgggt ggtccctct ctgcctcgca tctcgctagc tgccgcagca	7820
25	aggagttgca gcaaaaaacaa accaaacggt ccttaagatg atgcttagaa gataaaaactg	7880
	gtcacctaataa acttctccaa aagctgtgaa aatcttqcca tccaccaaga attgctgtaa	7940
	atgcgcaaaag agttagaaga aacaaacttg ctcagcaagc ccatgactca tccggcttcc	8000
30	actctctgag accagagcat gtgacttctg gcactggggc cgctggccca ccgaagtgtc	8060
	ctgtccggg gctgccaggt gctggggctt cagccgcctt ttgccccagg ctggacttca	8120
35	ggagccactg gaagggtggg aggtggaagg tgggagggc aatgcagggt tccccaccag	8180
	gaagctgtctt ttggctactt cccaggcccc aggttcaggg agtccctttt ccagatctat	8240
	tctccaaattt acatcctcac agggtccccct tagatatcac ccccaagcaag gcaulglgga	8300
40	ccgctgtggg tggctaccgg caggctgtag agcggcttag ccctgcaagg tgtaggaaa	8360
	ggctggggac ctgcaacacacc cactctctac ccacaccctt tctctctcag tgatgtctta	8420
	accctcgatca accctacaga ggacttaggt catggagata aattataataa gaaaggcaag	8480
45	caaaaacagt tttaaaaaat gtctttgaaa aatcttacct gtgttctgat gtgtcaaata	8540
	taagcatgca cttacagcca gcctgggtgt gtcatggccc agcccccgcctt gccccttcc	8600
50	gcaggctctg ttgccttgcctt ctgcatcata ttttcttccc ttccaggaag ccacattttt	8660
	agaaccctgg tqctqttcg gccrttcttc tcagggagca aatacatgaa ctggttcatt	8720
	catactgaca gtcttagtga gccagtgttt tacccctga gggatgtttt cattttgaaa	8780
55	gngaacctag gtacaylca atgcttaacc caaagaagcc aattcttacac taaatggtgg	8840
	gatttcaggc atggggagctt ctgattgtt taggggagag agaaggctla tgcccttgc	8900
60	uqaagctgtg ctgacttgcctt aagcttttctt ttctgtatata agaatgttcc tgaaaagcaa	8960
	tctataggtt ttgtttttat aaactcaatc ctactttaaa gacactaagt ttgccttcaagg	9020
	atatctgttc ttcttagggaa gtaaaaggctg atgataatta atattataac lltctgtatgac	9080
65	ttacaacgtg ccaggttctg ttctgaccat tgacccgcac ttttcttctc acactgactc	9140
	catgagggag gaatttattag tqlcttcgtt ttacaggtga ggaaaactgag gcacaaggag	9200

	gtcaagtaac tctteggtea cacaactagt tagtggcagg ttttaggattt gatccccatll	9260
	cactttacat tcccaataqt gctagactga agccttggga acccccaggct cccttaaaac	9320
5	cctgacacag ggcaacctcc cctctctttt catagg ggc ctt gaa gtt ttc tac Gly Leu Glu Val Phe Tyr 45	9374
10	cca gag ttq ggg aac att ggc tgc aag gtt gtt cct gat tgt aac aac Pro Glu Leu Gly Asn Ile Gly Cys Lys Val Val Pro Asp Cys Asn Asn 50 55 60 65	9422
15	tac aga cag aag atc acc tcc tgg atg gag ccg ata gtc aag ttc ccg Tyr Arg Gln Lys Ile Thr Ser Trp Met Glu Pro Ile Val Lys Phe Pro 70 75 80	9470
	ggg gcc gtg gac gtgagtgagc agggctgtgt taaaagcgggg ggacagttgg Gly Ala Val Asp 85	9522
20	gggtcttgat gttcctccag gcttgcctcat ctgggtqact tgaagctgcg tgcagagcca tcLggggttgg gctctggc tgcatggc tggccata aagttgtgcc ttcttttagag	9582
25	tggacttgag gagaaggctg ggggctggtg ggagggctgc agactggaaa ctttggtgat ggttgaggcc attctgcccc ggccttttcc tgctccccctg ctctgggagc tccttggctg 9762	9642
30	ccccctgagat ctgggaccccc aagctggtg gagaccagct ggtgtgtgt ttggagttgg cagataagtg gaagaaggcg tggtatqqqa ggagtcctct ctcaaggegc ggggtgcctg ggccttcaga tggcttctgg gctggagaga agctggatt taccaagaca gagcttcccc 9942	9702
35	agttgcctaa acaggcaqaq gtcagtgcattt cttgacgcattt tggatgcaca acatgagaca cgctcttctt ttgccccaaag ttgtgtaaaa gaagaacaaa gcatttttagg aactggaaat 10002	10062
40	aactatttta ctggcagtg attagtgcata ttacattttt ggtctggaa agcctgggaa agagtcatag aagtggaaaa tgggtctgtt gctgcctaa gcagccagag tggatgtttgg 10122	10182
45	ggaagctgtat ctggccctctt ggtctccgtt tctttatttc ttgactcctg ggccagactg agccccgcaga gtgtcaalgg tggagactgg cttttttttt gctgtgtttt agaaggaggt 10242	10302
50	gctgctgtca atctaggcag gacttggc acaaatgcctt ggagatggaa gggctgtggc cgggatgtga ccaaggcacac acaccaaggc tgaccttctg gtccttctg tgcaggglgg 10422	10482
55	cagggaaaggcc cacaaggggg ttcttcttc ccctcagtgc tgccttggtt aggccctcatg cagtcagctt agcccaagtag qacaacaagc tccagaggc ggcagcacac aagttactcc 10542	10602
60	aatgcgggaa alagactttt ctctttgaag tggatgttttac tatcagagag caaatagatt gtgatatggc tcgcaagtta tgaggctttt gtcattttttt gttttttttt tggatgtttt gcaaaacacga caaacacccctt ttcctttggg aagacgagca ataaagtgcacac 10662	10722
65	tgtgtgcaca tggatgttttac cacaacggc tgcacacccca tggatgttttac gtcagtcagccc tcttccctgg aaacccttgg actggggcaca cctgtggaca qccaaqagga gcaagagaaag 10782	10842
	gaaggcctgtt gattcttactt cctggatgg gaggtgcaga ggcggggcag gaacccaaatc tttttagagga gtatatgcata ctctgcggga caggacttc ggcacccatca tccatttact 10902	10962
	tgtcttgggg tttgtccctt atgtgtgtt ggtcagtcgg atgcatttttac aggcaggaa 11022	

cagtcggcaag ggaattctcc ccaaaggct cccttcccgcg tgctgatgtat gctgcagaca 11082
 5 tccactagat ggcaagttagt acttttctga gcacctggtg tcctccagag cagggatttg 11142
 aagcctgttg ggaaagtctt tggtttgcatt tggtttgcatt gccctggctt atcagcacat 11202
 tacacaaaga cgacatcacc cacgccacat cctgaagtct tgtactgtta aaccatcct 11262
 10 gcctcagccc qqqqqtqggg tggggagaaa acggatgtgg aacatagaga tttatgagga 11322
 acacagagct gtggtcaaga aagcctggga gagtgaagat ttacaaagg aagatttgat 11382
 15 cgaagctgtg ggaagaagaa aaacatagta ttctactact atgaaactat agtttatgga 11442
 aatgtAACAG acctatgaga aggcaGATTC agagagattt gccaaagtgg tcagtgagtg 11502
 ccaccttattt ccattcctgg gcactactcg gacctttcat ttaagtctcc aagttaaagta 11562
 20 ttctttcat caaqaactattt aatggcatgt gacaaaaagt caactcaaattt cggcttaagc 11622
 cagaaaaagag ggggaaacat tggttcacaa aactgaaaac tccaggata acttcaggcg 11682
 25 aggctggatc caggtgctca gtgtcaccag gaatctgtct cttgattgtt tttttctgtg 11742
 ctgttttcat ttgtaaatag agtctccct tggacaaaag atggccccaa cagccttagg 11802
 tttgtgtttt accaatttgtt caaccccaagt ggaaataggg tttccatcc attaaaagtc 11862
 30 acaggatgtgg ttcgttttgc tctttgtgcc tatccctgaa cctagcagtg accctctaca 11922
 gctccctta atctacaagg actgagagta gaggaggatg gttgcctga atgcaagtca 11982
 35 aggtgtgattt gatcagataa gggtgaaagg atgttaggtt ggaaatgcaa ggcaggctt 12042

 <210> 6
 <211> 2111
 <212> DNA
 40 <213> Homo Sapiens

 <220>
 <221> misc_feature
 <222> (1)..(2111)
 45 <223> public sequence NT_008300.6 (445091..447201 c)

 <220>
 <221> exon3
 <222> (1794)..(1892)
 50
 <220>
 <221> CDS
 <222> (1794)..(1892)

 55 <400> 6
 aatttccctt ttttgtatca tctttaatag ctatctctaa cagctatctc caatttctcg 60
 tctcccatct cagtcagecc tattccagcc aaatatcagt ctttctact ccactgaaac 120
 60 tggtcttatac gtgataacag tgatataccaa gttgtcaaat cagaltgtca catctgtgtc 180
 ctcatctggc cttaccttll agcagcattt gatatacgat atcaatttctt ctccctggag 240
 65 acactttcta tcctggctt tggacaccc ttttccagg tggccccca aactttttgg 300
 ctctaaagtc cccttccctt gattctcccc tcctctcaat tggatgtatgtt tttacagttc 360
 agccatgaga ccttttctt ttttgtgtc tgcctactgtt atttaaaggc catctctata 420

60394576 - 070818

	tcaatgacta ccaaatttat tcacccagtc ctgatctctc ctctgaatta cagaccgaca	480
	tgcataccta tattctgagg ggcatttcaa actcaagata tctaaaccag aactctgaat	540
5	gtccaccccttc aaacagcctc ctctccaagt cttatggta ccatttaatg gcccagttgc	600
	lcaagccaga aaccaggat tcatctcaa tttcttttc tccatcccc tccttactgc	660
10	cgtgccgttg aatctatcaq caaacccctgc cgattttacc taaaaatact cccccaaccc	720
	atctatggt ttctgtctcc atttctgcca tcctagccca agataccaac atcctccca	780
	tgtctactgct agtctcctaa ctggccctct ggccctctct ctggcctttg gagatcatc	840
15	attttctact cagcagccaa aagaatcctc taaaagcata aatcaqagca tatcactgctc	900
	ctgccccaaa ccctgaaaca gttctgtgtt gcactttgaa taaaacccca ctccttatecc	960
20	aggcttagag agccccctgt ggccctggct ctgtccctc ccctgcaccc gccatgtcca	1020
	ccctccccct tgcttgcctt tggccacaat ggtgtctttt ctqctcctc aaagctact	1080
	tttllccac ctgaagacct tcatacgagc tatttttttcc tttccagaac acttcccacc	1140
25	tctcatccca ccctgtttta ttttcatcat agcaactactc acacgttccc ttgttattt	1200
	atttattttt attatctgctg tcttttaccc agaacaagaa cttectgaaa gtatggacct	1260
30	tactgtcttg tgctctgttg ccclccggca cctagagcag ctctggcac acagtaggtg	1320
	ttcagcaaag gcttggtaaa gggatgaatg aatcctggtg ttctttgtga taggtaccgt	1380
	tcttgacaca aaataagtcc tcataatattt gttgaattga taaacaagtc actccaacta	1440
35	cactggattc ttcttgagag gaccagagac ctgttttagg ttcccttgat tcctccacgg	1500
	caclcagcac agcgctcaga acttcccagg tgcttggat gtatatcttgg ttqqattttat	1560
40	tcttcctcac aaagaggaag ccaagagtcc tgctgtctgc ttcacagatg ccatttctct	1620
	gagaataaaac gtatgtgag agggatgtca cgagagrag gataaactgg atgggacgcc	1680
	ataggggcta ctcagccatt ctgggttggaa atggcgggaa aaggaggctt gccccgttgt	1740
45	atttcctat ttctggccalg acgctacatt acttttcggc acatgtgtta cag ggc	1796
	Gly 1	
50	gca acc tat atc ctg gtg atg gtg gat cca gat gcc cct agc aga gca Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp Ala Pro Ser Arg Ala	1844
	5 10 15	
55	gaa ccc aga cag aga ttc tgg aga cat tgg ctg gta aca qat atc aag Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu Val Thr Asp Ile Lys	1892
	20 25 30	
60	gtaaaggc caagaggcac tctctgacgc atttgggtt gtatctggaa gattatcatt	1952
	caglgcagtt cccaccccttc atccaggaag gttggagcga ggacttctgt gtcccatctc	2012
	ccccagtgt cagctatcaa agccctcccc agtttccaa caatttcaaa aaglgcccc	2072
	9ggccagtgg tatgctggta acccagcttt ctagaagg	2111
65	<210> 7 <211> 10041 <212> DNA <213> Homo Sapiens	

```

<220>
<221> misc_feature
<222> (1)..(10041)
5 <223> public sequence NT_008300.6 (346211..356251 c)

<220>
<221> exon4
<222> (1408)..(1453)
10 <220>
<221> misc_feature
<222> (1423)..(1423)
<223> predicted SNP site
15 a is replaced by g in seqID 17 (176..176)
and in seqID 18 (115..115)

<220>
<221> exon5
20 <222> (3672)..(3785)

<220>
<221> CDS
<222> (3674)..(3784)
25 <400> 7
agactccaaa aaaaggaaag gaaggaagga aggaaggaag gaaggaagga aggaaggaag 60
gaaggaagaa agaaagaaag aaaggaagga agaaagaagc aagccagcaa ggaatgaagg 120
30 aaggaaggaa agatctatct cacctgtatc aggtgaagaa gacaaagtac aaggggctcg 180
tgggggaagg gtctatcctt ccacagtctc ctgcttcctt lccuaaccaa atgagaccta 240
35 gctctataagg aaccagcctt ggaagaggaa tatctgacca agaggccaa agtgaggaca 300
ataatagaca agctgcqtc taaaaggtag gagatcttgt gggttctgga ctgagcgagt 360
aggacttttag ttctataaaa ggttaaacctc aagctagggg attggatctc tctgcccc 420
40 gcccccacct caagcaggc ctccccacctc tcccccatcc tcccagcata ccatgagccc 480
ctcttctagg cttggtgaaa aggttagccag ggcacagttt ggtgtcggca qactcaatta 540
45 agagtacgc taatcacctt cattaaggct ataaaacccc acacctaggg qccttgggc 600
tccctggcat ggcctgggc ccctgagcca tggcacagac ccaggggccc ggctccagtc 660
50 tctttcatct gagcatctcc cagtcctgca caaacagttt ttagcacccca cccccgggg 720
cacctcatla lcccatagac actgagtgga ggggaggagg gagggaaaga qaggaggaa 780
actgaggctg tcccagtggc caaaqaaaact qccccagggt tggggctga gttaaactgag 840
55 ggtggggaga agctgccaca gaatcaggca gtgtagcctc ctctgtttt cccttcccc 900
gctggttgtct tgagccccat tgcttaatat tctaaaggct ccaaaggag cagagtccctc 960
60 cctccagaga agatgggctc cttaqtaaga tccagagaac tgcagacca agcgcctccc 1020
ttagatccga ccagaccctg caggtgtatgt ttagcccca gtctcaggca ccaataggga 1080
tcgtggaaaca gctctttggg aaglqlrcta cagggctctc ttagaggaa gglgtccatt 1140
65 tggccccctc agcaatacgt gtgtctgtcc tcttcctg gatccccctga ggtacccatt 1200
ccaaaacccc actcaclgcc ctgictucca aaagcagcca tacatcctcc tgctctccca 1260
rtttcctctt ctccccatctt acttcttctc ccatectctt tccctctccca lccclccct 1320

```

	gcagagagag aagcaatccc gcagagccca cgtggcctga agtgtgccc aqacataacc	1380
5	tacagccctg cttgtctt cttcaggc gcccacctga agaaaggaa gattcaggc	1440
	caggagtat caggtgagta agaccctagg gacccttcc cccagccagg gcatcagtga	1500
	cgatcacagg aatcaaaaacc tcacggtgct gcttgcgt gtgctagcac cagcctgaga	1560
10	gttttcttgg atgcccgtat ttcatctcca caactgcctt ttatggtcac tgccatcgt	1620
	gtccccattt taagaggata acactgaggc acagaccatc ctgcgccatg tgaatggcta	1680
	gtaaaagaatg cagccagctt cacagtgcag gggccatct gactggaggc tgcttgcac	1740
15	ccctgtccac ccccatcctt cccccagctc ccctctctca gatccctgcc ccaggaggat	1800
	gcccccaacc taccttggaaa gggatgtca ggtactggc tggccctctgg cacagcactg	1860
20	ggcacccaaac ttagccctgc agtccgtga agcagcgtgg gcttggcaca caagcaggcc	1920
	tggattcaaa tcccagcacc accactcaact gttggcctgg ctttaatctt caagcctcca	1980
	ggaaatggga agatggagga gatgtcataa gtgaaggcct acagcccagg gcctggcccg	2040
25	gagaagtgggt ctacgtatgt tgaacccttt cccctctcc tcacgcctca attctctgcc	2100
	tttggaggtt ttgaagaggg gattctttag gaaagcccta gtgagaggca ttagtctcta	2160
30	tctcccttgc ggagatagag gataaattct ccattgcaca aatccagtga gttctcaett	2220
	cagacttgc cgcacagaac tgaggtcacc agttttggc agcgtcaagt gtgcccacgg	2280
	tggagtgtac tccacagccc acatgcccac acatcttgcac cacactgact tgcccatct	2340
35	gcctccata cctaagcaca cgaggagact ccctggcacc cacctgtctt cagaagtctt	2400
	tctcagccct ggcgggtttt aggaagaaat gtccttac clgcattac acacacacag	2460
40	qcctccaaa acagcatcg ggcgtttgca cactggctgt tcacgtacac acacccaccc	2520
	cgagtggac ctccaaaccac acacatggct ctcagcacac ccagctctgc tctcaactctg	2580
	agtgcacatg cgcgtgcaca cacacacaca cacacacaca cacacacata tatctgtatct	2640
45	gcgatccaaa aatgcaggaa aagatgaatc gtcaggaatt ccccggttgt tggcatgtg	2700
	gaagccacca gtggccggaa gtgcacagtg agggactctt ccagtgtcc aagcagcaag	2760
50	tctgtctttt gtaatgtat caggagagga aagaaggaa gatggccata aaggacaaat	2820
	actccctccc accaacatec tcattgcataa tgcttcttc ccacccaccc agaaagattc	2880
	ttggctccag cctggctctc ctccctggaa aggaagatgg agagcggaa ggggatgtgg	2940
55	aaagggtgaa gcaclctggg aaagaggccc accttgaag acctaggcca ggccagcagg	3000
	accttgaactc tgcatggact tqgggaaaga aigaggatg tccccacccc actttttttt	3060
60	ccccaaaaaa agagctcagc ctccctggagg gagagagaag gactgggaga gggacttct	3120
	ccctgtatcc cccgacccaccc agtccttaca gtcctccatga tcaaaccctt gaacccagtc	3180
	ccacccaccc tcgttaaccac ttccaaagaccc tccagacqca qccaagaaat tggcttctt	3240
65	ggggallqag ttcataagca aatgccggaa aaataaataa aaatctctg ccaacacac	3300
	tcacagagga agatggqctt ctcttggcaggc aggaaalcca tcaggataga ctaaagcaaa	3360

60394576 . 070802

	aacaggacaa tgccaacacc cactggccca laaaggcctg ggagccacgc cagcacccaa	3420
	ggccttcaag ccagctcagg gtgttcctc ttggaccacc tcctccccca ctctcacccc	3480
5	cacccaaaata tgcattttttt aaatgaaaat actgaagctg agatcccgcc aaatacaaata	3540
	cctaattgttt tgcaggggcg agtgcgtgate tggggcccca gggaggctag caggaccatg	3600
10	ggaaggcaga gagctctgcc ctgcgttagta tggatgttg atgcacactt ctctctgtct	3660
	tctacctata gcc tac cag gct ccc tcc cca ccg gca cac agt ggc ttc	3709
	Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe	
	1 5 10	
15	cat cgc tac cag ttc ttt gtc tat ctt cag gaa gga aaa gtc atc tct	3757
	His Arg Tyr Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser	
	15 20 25	
20	ctc ctt ccc aag gaa aac aaa act cga ggtaagacat cccatcctt	3804
	Leu Leu Pro Lys Glu Asn Lys Thr Arg	
	30 35	
	ccagccacca cgggggtgaa ttctgggtg gaacaaacgc ttcaaggcatt tacggggaaa	3864
25	ttggggcctt gaccccaact catgagcctt ggagactggg gtcgaaggct ggtgtttttt	3924
	gggagtgaac tggaggagt gcccctaaaga tggatggatc aggtgtggc tcccttgaga	3984
30	agtaagtcac acagaccagt gtagaggctc cagttgggtg gyltggtatg taaggcaggg	4044
	tctcaaagca ggtgactgca gccaggaga agatcccc gaaactttqc atggcatgtg	4104
	caggcagctg cccctgaact ttggcaagtc attgtgtctt gtaattgagt ctgatctggg	4164
35	caagaggccc ctggaaalgc accctgcccag tctatcctca tgggttgcta attactgggt	4224
	cggctgtgcc atctcagtga tgagggaaatg atgtccccca ccattttttt tccactagct	4284
40	tctctgagca gccccttgcgacttcttggag cagaccaact ccctatagcc tccataatggc	4344
	gaacccccc tgcttggcct cggctctctt gcccaccaggc ccaggagagca ggagcggtgg	4404
	ccacttggcc actgcagcat gccacaagtg tctcacctga cccagtgtga gccatgcccc	4464
45	cgactttttttt acccgccatc tgagccttag atttctttt cagaaaatgg gaaataaaagc	4524
	cacccctgca ctctcagctc cccaaattttt gggaaaggaa gacaatagag actgaacaat	4584
50	qccatctgga agtgcagaa gacatcccc caggctggca ggaaggcaag tggagggttg	4644
	ggtgtcgtca gccalltctt ctcacccagt gacctcgacc aaacatccac ccactaggaa	4704
	gctggcacag gtattgtac aaccgcgtatc cccggaggccc cacaggagag gaacatggag	4764
55	cccagaaagg cacaggttga ggcacaggcc cccctcagatg tgccgagcag ggcacccatca	4824
	tcacagaggg ctccaggcata cccctggagt ggagctggtg caaaacttacc accagtgaac	4884
60	cttccccagg tgccctcagc accttgcctt tccgtgaatc tggatggatc gctttgcctt	4944
	agacatactl llaagtttag atgcagagag aaggaaacaa aatlatgtat cttggccaga	5004
	ccaaagaaaa ctgtctccag caccctggct accgatctca ccaaagatgt cqqgacrtca	5064
65	gcctatqaca ttctcccacc catgggttcc tctataaaatg aagcagtgc tctgagcccc	5124
	accatgagac tccgggatca ccacgcctat gcccaggccg ggagtcttc ggggtgaay	5184
	tgctttcgtga aggccaggatg gagggaaagga tagacatcg ggagatgcca tgaggaggga	5244

	cagacaaggg agaggaaggg agctgcatac accaaagacg ctccccact ctgcgtgctg	5304
5	gggtcttga ggctcgacca cagccctgtcc tccagaaacc acagtggct gagggaccag	5364
	cagaggcctg acccacccct cccaaagatgt tggtgaccag gcaaggaaagc caaaacgcgc	5424
	tgggcctgaa acgcgttgtt tggcttgcgt tactttccca cccccccct tgaccaccc	5484
10	tttgcctatg gggaaacttag aaactggact tgcacatgtc cagggttaa aatggccctc	5544
	ccaggctggg cgccgtggct cacacccata atccccagcac tttggagac tgagggtggc	5604
15	ggatcacgag gtcaggagat tgagaccatc ctggctaaca tggtaaacc ctgtctctac	5664
	taaaaataca aaaaaaaaaaa aaaattagcc cagcgtggcg gtgggcgcct atagtcccag	5724
	ctactccgga ggctgaggca ggagaacagc algaacctgg gaggccgagc ttgttagtgag	5784
20	ctgagatcgc gccactacac tccagcctgg ggacagagag aqactccgtc tcaaaaaaaaac	5844
	aaaaaaaaaaa aaatggccct cccagatgaa agccccagcc aaggcccacg tcttgccatg	5904
-25	lgatgccttc catgtggcca gcgtgtgtc gggaaacttca gaggagacac aatcaaggc	5964
	accaggaatt ttccctgtcc ccaagtggct tacagtctac tttggaaaca attaaacaac	6024
	aatgcaaagc aggtttaat taaacactla altgtgttat tctaactgtt aggacaattg	6084
30	gaaglctgtg gactcagggta tagatcttt cccttctgtt cataccttc cctccagaac	6144
	actgtgtct ctccagccag agaggatata aagggtttt attqtagagt gaggcaatct	6204
35	gactccct ctctctagtt agtagttca ctltgaggcc aatcaatgaa cctcgctggg	6264
	cctcagttt ctaatctaga aaataagggg ttggctccag atqatctaca aatttctctg	6324
	gccctgaggg catggggagg aggttggagc agggctggc cctagagaag gaggagatct	6384
40	gaatccttga gtgaaggaag agggagctt gggttggaga aaccatagat ttttttttt	6444
	taaqtgagaa aatggagggg aagggaaaaa acaaggcttct cagaagtccg gcagaggagt	6504
45	gcttttgcgaa ggttgcactt gggaaagccgt tagaaattat ttaccagaga agcaacactc	6564
	aaaaglcact ctggagac gacaaacctg gaqgggcagg cagaactgac cagaatgaga	6624
	ggacaggaca ccaccaqaqg ggtcgagaca tgaggcagt aggtgtggat gaaccagg	6684
50	gcgattgc当地 aagcagaaac cccaaaggctt tgcacgtatgt tgaaacactc agggcctgg	6744
	gctctctgna gcacctggac ggtgcctggg gacttacgga gcccaggct ggaccagtgc	6804
55	ttgggaggcc agagcatggg qtccctgcca tgaacacagg tccctaggact tgccttgc	6864
	tagatttctc tccaggggag agaatccgtg gaggaageag yylqlcaggg gtgggagg	6924
	gggcagtgtq gtataaccag ctgcttgtt cccctgggtg aacaggcagg aaaagcttct	6984
60	agagactcga gggaaaggt cagccatcag gagccaggtt gaggggcqtc tatggcagac	7044
	ggccactga aaaaaatccc gttcaggccctt tggatcttata gggtaaatga	7104
65	caggcttagg gtgaqaqq aggcttgaga tgattctagt agagctgtc atgactatgt	7164
	agggggatgt ggggatggca ttgttcagac ccaaggaggl qaaccctcc ccaggattcc	7224
	tctttttact gtgaccacag gacaatgtac atgtggtcag caaatgagat ctgtatgtcc	7284

	cgactctgat	cgtttgacct	tggccaggtg	accccccctc	cctgggtctc	aattttgacc	7344
	tcatcaagat	acagggttag	actagttgat	ctgcaattga	atgacctgaa	attgatttgc	7404
5	tcccaagcttc	ccggaaatgtc	ctggcctcccc	tgttctgcaca	aataaatgca	gggacccaact	7464
	gatgaggctc	gactttccca	gccagccctg	atacttccta	gtagaggctc	tgcacatcagac	7524
10	cccacqaqqc	caggagcaca	gcaagggtcg	tggaggggct	tagactacc	ccacttctag	7584
	ctgatgaaac	cctgcccgtgc	caccccaactc	ctccctacccct	tgcacagctc	ctccctctct	7644
	ggccacaclltc	aaaccacgccc	tggcataact	cagtaatgcc	ttctgcattc	gctcagctga	7704
15	tgcagatgc	ctagtgtata	cttggcgctg	ttttqggqtc	tgttaagaagt	caccatgtga	7764
	tgggtgtccct	gctctcatgc	atgtgtggag	tcaccctgg	gtagctctat	ggataggggg	7824
20	tgctcaactga	gagagtcaagg	ggagatattc	aaacagcaca	agcataactg	gtccccaaaga	7884
	taggaggaac	agagagaccc	ctgaaatcta	aacagcctta	actgggcct	ctgacaaggg	7944
	ctggccggcc	agcgtcacag	gacagctatg	tgacgcttgt	cccaaaggcc	acactttgcc	8004
25	tttccagatt	tgcagatgcc	attatctcag	gcctgtgaca	tgcctggctg	tctccctctg	8064
	attcttaagt	aacacccttc	tgagggggac	atctttcctt	ccaggcagat	ttcatttgca	8124
	ccttgtacca	tlgcccattgt	gtgcccaca	ctgcccagacg	tccctgaact	acttctttgc	8184
30	tcacatctgc	ctgccccacca	ctgcccacatc	gggcgtcica	ttctccctt	ccagaccccg	8244
	caatgccttg	cccaacttcc	aagcaccagt	cccgctgtgt	ggccctccat	cacatccct	8304
35	agaacaaccc	cacaaggctg	atccgctgag	ccaggctcc	cccaacctgc	tgttggcatac	8364
	tctagtctct	atcgctttct	ccactcacac	agcatttgtt	gtgcctcccc	agacatctgc	8424
	tlgatctct	tttgtatttc	taaatcactt	gccactcgtg	ggggcccccc	ttccaccatg	8484
40	agactatcag	ttccttqaga	qaagcgtca	cctctgcctt	ctttcccttc	tcccacctgt	8544
	ccctgtgccc	agcccactgt	tggctttcag	ttaatqagt	tgtctgtgt	tgagacacac	8604
45	ccacatatgc	acalccacaaa	gcatttcaca	cacatggatc	tgcacgttct	cacaccccca	8664
	catttgcaaa	cacacgtcaq	cttgcctcaq	cacctqagaa	ggcagccacc	tctgtgagta	8724
	ccatgcccac	aagtgacagt	ctctgcattt	atcgacgtcc	tccctgtccc	aatcccaggc	8784
50	aggcagggt	gcctctccac	cagcccttc	accagccacc	ctggccctcc	ctggcccttcc	8844
	catttgccacc	catgtcaagg	gggcaggaca	gtggccatca	cccttacccc	agggcaaacg	8904
55	agggaaagcaa	agcctgaaac	agtgagttgc	atgacttacc	aaggcgtcc	acagtttaaca	8964
	gccatttgac	tttcttctqa	cttccaggga	agggctgttt	gcacctgcca	ccgccttgc	9024
	tacatccccg	ctctctgac	ccagacttct	gggccccagt	ggggatacag	acagagcaca	9084
60	gttgcatact	gtgtccccag	gcctccitgg	ccagcagcag	ctgccccagaa	tggtccacgc	9144
	acacacccccc	tccctggctg	tggctgagcc	tcagaatct	ggggagctgc	ggtctgtgcc	9204
65	acacccacacc	tgggtcacct	cacccagagg	tgtcacaagc	cacccraca	ctccctctgc	9264
	aagcgagcag	gggaaatccc	ctctctgccc	cacgagagga	ratgcccacag	cttgcctccag	9324
	aagccactcc	aaggccgagc	gcccaaggatg	gtgtcaacgt	gctggaaact	aacqggcaag	9384

```

agaggcttgg gtcaccaccc ccagggcagc ctcagggccc tcagagggga aacttgccta 9444
ctgcaaaatg ggagtccctgg gccaggcagc cataagccgc tccctagcac cacaggcacc 9504
5 caaattcggc agtcagagg tggAACCCACCA caccCACAGG ggtcaagAGAG gagAGACGGG 9564
cagggacLLg ccacatAGGG ggttGAAGTC ttgtttGTGG catCTGAAGG catGCAGGCC 9624
10 tcggtggttg agaaAGAAGG ggaATCATGG ggccaAGGGC agcAGAGCCC cacqcaqctt 9684
caagAGTAGG agatGTTGCT gcAGAGGGAT gaggctGGGA ggAAATCCAC cAGCCTAGTA 9744
15 gcagaACACA gcAGCACATG gtgtcAGGAA gcatGGAGGG aAGAGGCTG agaggAAAGG 9804
gggcctgcC tggAGACCAG CCCCCGCCCC CAGACTGCC CCCCACAGCG cacACACAGA 9864
cAGACACAAA cacACTCAGC CCACCCCTGT gacCTCCAT ggtggTTCAC ctggcccAGC 9924
20 CCCAGCATCC CCCAGGGCCC AGCCCTGAT cactgtgtgg gtttgtggca ggaggcgctc 9984
ttcctccaaa gggccccacc cacctggccg agcgggtaga actgggggcc tggcagc 10041

25 <210> 8
<211> 6241
<212> DNA
<213> Homo Sapiens

30 <220>
<221> misc_feature
<222> (1)..(6241)
<223> public sequence NT_008300.6 (339931..346171 c)

35 <220>
<221> exon6
<222> (5010)..(5278)

40 <220>
<221> CDS
<222> (5012)..(5164)

45 <220>
<221> misc_feature
<222> (5150)..(5150)
<223> insertion-deletion site
      (insertion of a nucleic acid 'a' gives seqID 16)

50 <220>
<221> 3'UTR
<222> (5165)..(5278)
<223> partial

55 <220>
<221> polyA_signal
<222> (5273)..(5278)

60 <400> 8
aggcACCCAG gctggGAact gcctgAAATC tccatgAgCC ccaggcggAC ctccTCAAGA 60
65 agcaAGTCCC agtccAGCCG AACATCAGCC atccCAAGCA atgtGAACA catcgtGAAG 120
aaACCACACA gataAGCCCC caqATTCTAC cccAAACTCC agccaATTGL tgccACTACC 180
70 aaATATTGGA atttatCAct ggaAGCTAAA ctggcAGCTT gaagtCTCCC atcCTCCTGA 240
tccCTTTCTT gccccAAACC CAGCTGAGA agggAGGAGG aggCCTAGGG gatgcACTGG 300
75 ctggATAATT agaACCAGCT tccACTCTTC agccCTCCTG ggtcCTCTCT ccAGAGGAGA 360

```

	cgcaggtcac	ccctcaccac	cccggaaactc	tgctggcata	ccctccctct	cactcctgcc	420
5	aaagtcaaag	tgcggccac	tacccgaaag	cctctggagc	ccgcagtata	agccactcag	480
	caaatctcct	cccgaggata	ttggatgtga	tggaacgagg	gcttttcagg	tcctgagtgg	540
	aagtctctgg	atgtcattta	atggcacaaa	aatanaagt	ctcaggcctc	ccgcctacca	600
10	ccgcgggac	atcatataatc	tcggcagaaa	ggtatcagtg	cttctatcaa	tgtggccggt	660
	cccggtacga	gccagacaaa	ttacagacct	yagccggcgt	ctctgtgggg	cggttattag	720
15	gctggagttg	agtttttgtt	ttttctcttc	agggccacca	tcccccaaca	agaagcagtg	780
	gtgccttac	ctaagaggag	aggcagcag	gccaggaaag	agggtgatca	cctgggttag	840
	atacagatca	ccacctgaag	gaacagggt	gttcggagcc	ggagccagaa	ccagtggccc	900
20	ctgtccttac	agagaccgc	ctggacaag	aggctgcgc	agctccaacc	tcaaggcttc	960
	agcaacccac	tccttgtct	ccagccacc	tgaacggggg	cccatggag	aatggagctg	1020
25	ggylgggggg	agcctccctc	cctcagettc	ctaccccaact	tcccaggtca	agagggaaat	1080
	acatcaggat	ggaaatacat	caggacggaa	atactgagga	ccaaggacca	gtcaagggtgg	1140
	aaacgtccct	tggctacage	agggtcaccc	aggaactgct	gtcctcggt	gttggaaagt	1200
30	cttaagagtg	ggctgtcta	ggtgcagtga	ctcacacccg	taatcccagc	actttggag	1260
	gcccaggcag	gaggatca	tggccagga	gtttgagatc	agcctggaca	acatagggag	1320
35	accctgtctc	tacagataat	tttttttaa	ttacccaggt	gtggtagcac	acaactgtgg	1380
	ttccaqctac	tcaggaggct	gaggcaagag	gatcacctga	gcctgggagg	tcaaggctgt	1440
	aatgagaggt	gatagtqcca	ctgcacccca	gcctggagaa	tggctctgcc	gaagggaaatc	1500
40	caggggtgtc	actaggggtg	gggaccttgc	ctgcgtggga	cccagactac	aaagttctac	1560
	tgctggccaa	ggccacccctt	tgaaggacca	gactgtgatt	aggttagtat	tacttgtatt	1620
45	atattagaaa	ctcgccaagg	gcaggggta	aaataaggca	gctccgctga	ccccatgcct	1680
	ggcagaggac	cagggtgatg	gctgccagtg	ttaaaggggc	cctgggtct	gacctcatct	1740
	aagagcttgc	acttctggca	ggggagccga	tccccatcc	taccaagtcc	tgcaaggggg	1800
50	ccagggtaag	ctggatgtgg	cagagcacct	ctcagtcct	cagatgcctg	cttagcccta	1860
	actggccaa	tggagccgtc	cagagttgt	ccagcacgtg	acaaaccact	cccaaacagt	1920
55	ggaaagccaa	ccccacagat	ttggaaagct	tgggtggag	agagtgactg	tcaagacctg	1980
	ccccaggggcc	ttttccaaac	atccctgcca	tcatcacagt	ccctggact	aaagctccga	2040
	ggaggagccc	cttccccac	tgcgtgaagac	aggggacatc	tttccccct	agaggctaat	2100
60	aggaggccag	gcacagagtg	gtggcccaga	aatccccag	tggctttaga	taaggttatac	2160
	aggctctccc	cttggccccc	atctgaaaca	tcgctgaaat	ctaacatctc	tcatcactaa	2220
65	catctttct	atcagcactq	accagtaagt	gagagtgc	ggactgc	lqc atctggctat	2280
	gcaggatgag	aacagcaccc	cctacggttg	tgcagggtal	aacctgtaca	tcatatgtgg	2340
	cagccctgtc	ctgagatgcc	caggcctctg	tgtcccagcc	acagggcagg	tggctcatga	2400

	gacaggcata agacagcttc ctccctttt ctcctttag ccaaggcctc ccaccaccaa	2460
	ggcctctqaa gtccccacac tgccccatacc ctcgtccccag gaagagcctg gccttaggtgc	2520
5	atccaggata cagtggagtg gcccaagcac agacagcagc gctccctttt tcctgcctgg	2580
	cccgccgactc tctgcctctc gggcacagca gtccataca ggaaggggct ttccctctgag	2640
	cgggtgaccac agaagccgca gcctcctagc tctccagetc tccccacaca ccccaaggc	2700
10	ctgactcacc ccctctgcct tctagaactt gcccctctgtg gcagggaaaca ggggggtgtaa	2760
	ggagtgcgg ccaagccagg aaaggaggca gcaagaggct agccggcacy gggactgggt	2820
15	tccaggagct ggccaggggag cccgagaacc acagggttgt cgacttggga ctggaaagga	2880
	accacgagtg ccccaggca ctgaaggctc agaggcttcc gaactattga ggggagccag	2940
20	ctctcgctct ggactcaggg gatttcttcc cccgagcttg ccccaaggctc gtatgggtt	3000
	tcaattccct tccctgcgtg caggttccctt cctgccttac ggtggggaaag gcccccttgt	3060
	ccctgttccct gccacacacc aattcaccca taagccctta ttaccctggc atgtgcgcgt	3120
25	gcgcacacac acgtgcaccc acacatacac atgtgtgtgc acacactcg tcccaagggtc	3180
	ataataccat aattcccttc cttccagcca gctgggttcc tagggtgcaag aggctcaggc	3240
30	acagccccctc tgaaaggaca ctcaggglgac actgagcctc ctaggagaac cccgctgcag	3300
	tcaqtacacc tccagccggaa gaggtttctt cagcaccttg qccacccctt ctgc当地aga	3360
	actcctggaa ccccttggaaa tccagcccccc ctggggccca tctgagcatg cagtcactga	3420
35	gtcgaatagg atttggggaa atctgcaaag gccttcccc atccctgcca ctacgaagct	3480
	tccttgcag tacccagct tgccttcag tcatcccttc tgcattgg gaatgggaaac	3540
40	aglcculcccl ttttggatt ggtgtgagga ttaagcgaat tatgtgtcaa gctcttacag	3600
	tagtaccaac acatagggtgg qgtgttactg ctattatgt ttgcattatt gttattgtta	3660
	ttttgggttt cagattaaaa attatccctt ctttggaaac atctttgaaa tgcagcagc	3720
45	ctctgggagg cagcaacgc lcaatgggt tgcatactta attgctatga aaagggtgg	3780
	gggtgctgac cgaggagaca ggttaggtgg agggtgggtg gcccgtqagt ctcaaqgtgg	3840
50	gttctgagag ctcccaaccc ctgcctct tctgtctggg tattcagaggt gactggctc	3900
	cattaatcca ctgtgtatcg accactgcct agctctgc tggacagggl cltccccagg	3960
	cactctctt ccacccctggg aggttgcgc atttagaaaa acaacaacaa caacaaaaag	4020
55	taggalgeccc agtaaaaattt gaattccaga taaacaacgg aaaaaaaaaat tcatgtaaagg	4080
	atgctccatg cagaatgtgg gacacgtca caccaacaag cgatttgcta tttatclgag	4140
60	alcaaaattha aactacttcc tgcctttcc tggcaagcat ctgcctcagg tgcctqacat	4200
	ccagggtttg gaaacaaga acgtgggaal algcclctga ccactccct cagcctcact	4260
	ccatgccccca ggctgtgccc ttctcgccctt gccaacaggc ctgttacac ttgccccaga	4320
65	agatqcgctt ccctgtacac acttgggtttt tctcccttg aggaggaacc cctccgaccc	4380
	gggcaccgtc tgggtttctg caggcccaqr tcccccggca ggtccaggg gctyccccclc	4440
	caclclgccc atggggaaagg cacactggtg cacccctgt cctggacag agacqagqaq	4500

	gtgtttgggg ttttaggggag aayatgtaaa gaagagcaca gatacttcca ccaatctcag	4560
5	ccccagcagc aggggtgagc aggtgtgaag gaaaqtagcg ggtgaaggct cctgtccctgg	4620
	ggtcagggga gcctgcaggg agagggcatt agtagagggaa gtggcagcc cgcttgactg	4680
10	ggcttcctt ctgcccggga gcgagaaggc agtttgcata dacagaggtg ctggggagg	4740
	agattcagac attctcaact ggggctggc ttcagtccc cattgccctc tggggagcag	4800
	gcagctctgc tgctgltcag aggctctgct aacatctggg ggctacctct gcaacaggag	4860
15	ccggaaagctc ccccagccac tacgccccat ggccatcgca tcacctgctg cctcggtctc	4920
	cacgagccgg ctcccccttc aggcctctgc cagcttcaa gtccatggct ctttctcaact	4980
	gcctccctgt ggtctgtttg tccttcagg c tct tgg aaa atg gac aga ttt	5032
20	Scr Trp Lys Met Asp Arg Phe 1 5	
	ctg aac cgc ttc cac ctg ggc gaa cct gaa gca agc acc cag llc atg	5080
	Leu Asn Arg Phe His Leu Gly Glu Pro Glu Ala Ser Thr Gln Phe Met	
	10 15 20	
25	acc cag aac tac cag gac tca cca acc ctc cag gct ccc aga gga agg	5128
	Thr Gln Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala Pro Arg Gly Arg	
	25 30 35	
30	gcc agc gag ccc aag cac aaa acc agg cgg aga tag ctgcctgcta	5174
	Ala Ser Glu Pro Lys His Lys Thr Arg Arg Arg	
	40 45 50	
35	gatagccggc ttgcacatcc gggcallytgg ccacactgct caccacccac gatgtggta	5234
	tggaaacccccc tctggataca gaacccttc tttccaaat taaaaaaaaa aatcatccag	5294
	ggcttgggtgc ttgttatcct agcttgtatt cctttgaaaa gcaacaggag gaaattggcc	5354
40	aatatccagc actggggact gaggcttcc tglaaatgca aggggttggg tggggagtgg	5414
	ggqqaataga cggacaggca ggatatagtg atccctacgc ctgaactgtc actgagtcgc	5474
	tgtgagacct cacggaaacc lcltaccctc cccaaagttc tgctccctt gttgtgacac	5534
45	ggaggttagga agaagtggga gagggtggg ttatccatcac cacgggcctt tccagctctg	5594
	acatccqaca acgttcattc agcgcattt gttgcattgtg ggggcacact ggccttcctc	5654
50	caggcaactgg agatgcattc ctgccllcattc cccatgcattc ggtagatagg gcagacacag	5714
	catggtgctc acctcatggc gttccagtc cagcaqggga agcaaatatc cccaaacaaat	5774
55	gataaaactgt gataagtatg atgagggtgc aaagggggccc tgatttctga cggagagaaa	5834
	gggttgaaga ctcaggcaga gccctgatgt ttaactlca ggtgtaaaag tggcttaggg	5894
	gttaccaggc gatgtggagac aggaatagag aacattccaa agctcctccg agcccctgtg	5954
60	gctggcaggg gcatgcattt catcllccat gaaatgaaag gccatgcattg tggclgeccgc	6014
	aaaylatgaa agggcaagaa cagtatgacg qgtggcataaa ggccagacca tgccaggcct	6074
	tlgttgccat gttaaagctc tggattgtat ccagaaagca gtggggaaatc attaaatgt	6134
65	aaatgatattt ggccaggcat ggtggctcgt gcctataalc ccagtgcattt ggqaggccaa	6194
	ggccagagga tgcgttgaag ccaaqagttc aagaccagcc tggggcag	6241

```

<210> 9
<211> 14334
<212> DNA
5 <213> Homo Sapiens

<220>
<221> misc_feature
<222> (1)..(14334)
10 <223> public sequence AC105046.2 (20000..34333 c)

<220>
<221> exon5
<222> (1023)..(1136)
15 <220>
<221> CDS
<222> (1025)..(1135)

20 <220>
<221> exon6
<222> (12443)..(12711)
<223> seqID 1 and seqID 2

25 <220>
<221> CDS
<222> (12445)..(12609)

<220>
30 <221> 3'UTR
<222> (12610)..(12711)
<223> partial

<220>
35 <221> polyA signal
<222> (12706)..(12711)

<400> 9
40 aattgtataa cgactcacta lagggcgaat tccccggttt ttggcatgct ggaagccacc 60
      aglggccgga agtgcacagt gagggactct tcccaqtgctc caagcagcaa gtcgtgttt 120
      tgtttaatga tcaqqagagg aaagaagggg agatggccat aaaggaccaa tactccctcc 180
      caccaacate ctcatcgaa atgtttttt cccacccacc cagaaagatt cttggctcca 240
      gcctggctct ctccttgga aaggaagatg gagagcggga agggqagtgg gaaggggtga 300
      agcaactctgg gaaagaggcc caccttgaaa gaccttaggcc aggccagcag gaccttgact 360
      ctgcatggac ttggggaaag aatgagggat gtcccccaccc caclcccccc gccccaaaaaa 420
      aagagctcag ctccttgag ggagagagaa ggactggag agggaaacttc tccctqatgc 480
      cccgagcacc cagctctac agctcccaty atcaaacctt tgaacccagt cccacccacc 540
      rtgcgtaaacc cllccagacc ctccagacgc agccaagaaa ttggcttctt tggggattga 600
      gttcataagc aaatqccgaa aaaataaata aaaatccctt gccaaacaca ctcacagagg 660
      aagatggct tctcttgagg cagggaaatcc atcagytag actaaagcaa aaacaggaca 720
      atqccaaacac ccactggccc ataaaggcct gggagccacg ccagcaccca aggccttcaa 780
      gccagctcag ggtgtttctt ctggaccac etctccccc actctcaccce ccacccaaat 840
      atgtacatct taaatgaaaa tactgaagct gagatcccyc caaatacaaa tccataatgll 900
      ttgcaggggcc gaqlqctgat ctggggccccc agggaggtca gcaggaccat gggagggcag 960

```

	agagctctgc cctcgglagt atggatgtt gatgctcacc tctctccgtc ttctacctat	1020
5	agcc tac cag gct ccc tcc cca ccg gca cac agt ggc ttc cat cgc tac Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe His Arg Tyr 1 5 10 15	1069
10	cag ttc ttt gtc lat cll cag gaa gga aaa gtc atc tct ctc ctt ccc Gln Phe Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser Leu Leu Pro 20 25 30	1117
	aag gaa aac aaa act cga ggtaagacct ccccatcctt ccagccacca Lys Glu Asn Lys Thr Arg 35	1165
15	ccgggtgaa ttctgggtg gaacaaaacgc ttcaaggcatt tacggggaaa ttggggcett gaccccccact catgagcctt ggagactggg gtcgaaggct ggtgtttctt gggagtgaac	1225 1285
20	tggaggagt gcccctaaga tggggcatac aggtgtggtc tcccttgaga agtaagtccc acagaccagt gtagaggcgc cagttgggtg ggttggatg taaggcaggg tctcaaagca	1345 1405
25	ggtgactgca gccaggaga agaglaccgc gaaactttgc atggcatgtg caggcagctg ccccctgaact ttggcaagtc attgtgcctt gtaattgagt ctgatctggg caagagggcc	1465 1525
	ctggaatgcc accctgccag tctatcctca tgggttgcta attactgggt cggctgtgcc	1585
30	atctcagtga tqaqggqaag atgtccccca ccatcccccc tccactagct tctctgagca gcccttccttgc gcttctggag cagaccaact ccctatacgcc tcctaattggc gaaccccgcc	1645 1705
35	tgcttggcct cggtcctctt gecaccaggc ccagggagca ggagcggtgg ccacttggcc actgcagcat gccacaagtg tctcacctga cccagtgtga gccatgcccc cgactcccc	1765 1825
	acccgccttc tgagccttag atttctcttgc cagaaaatgg gaaataaagc cacccttgca	1885
40	ctctcagtc cccaaattttga gyaaggaa gacaatagag actgaacaat gccatctgg agtgcagaa gcacatcccc caggctggca ggaaggcaaq tgggagggtg ggtgtcgta	1945 2005
45	gccatttcctt ctcacccagt gacctcgacc aaacatccac ccactaggaa gctggcacag gtattgtcac aaccgcgtatc cgggaggccc cacaggagag gaacatggag cccagaaagg	2065 2125
	cacaggttga gcgacagggc ctttcagatc tgccgagcag ggcccaccta tcacagagg 50 ctccaggcaca ccctggagt ggagctggc caaacttacc accagtgaac ttccccagg	2185 2245
	tgccttcagc accttgccaa tccgtgaatc tgtttaaacg gctttgccaa agacatactt ttaagtttag atgcagagag aaggaaacaa aattatgtat cttggccaga ccaagaaaaa	2305 2365
55	ctgtctccat caccctggct accgatcua ccaaagatgt cgggacctca gccttatgaca ttctccaccatgggttc tcataaaatg aagcagtgc tctgagcccc accatgagac	2425 2485
60	tccggalca ccacgttat gcccaggcgc ggagtcttc ggggtgaag tgctttctga aggcaggatc gagggaaagga tagagcatac ggagatgc tgaggaggaa cagacaagg agaggaaggag agctgcatac accaaagacg ctccctact ctgcgtgtc gggctctga	2545 2605 2665
65	ggctcgacca cayctglcc tcaagaaacc acagtggct gagggaccag cagaggcctg acccacccctt ccaagatgt tggtgaccag gcaaggaagc caaaacgcgc tggccctgaa	2725 2785

	acgcgttgtt tggctttgtg tacttccccca cccccccctc tgaccacactt ttgtccatg	2845
	ggaaacttag aaactggact tgcatactgtt cagggtttaa aatggccctc ccaggctggg	2905
5	cgcgggtggct cacacccata atccccagcac tttgggagac tgaggtgggc ggatcacgag	2965
	gtcaggagat tgagaccate ctggcataaca tggtaaaacc ctgtctctac taaaaataaca	3025
10	aaaaaaaaaaa aaaattagcc cagcggtggcg gtggccqccc atagtcccag ctactccgga	3085
	ggctgaqqca ggagaacacgc atgaacctgg gaggcggagc ttgttagtgag ctgagatcgc	3145
	gccactacac tccagcctgg ggacagagag agactccgtc tcaaaaaaaaaac aaaaaaaaaaa	3205
15	aaatggccct cccagatgaa agccccagcc aaggcccac tcttgccatg tgatgccttc	3265
	catgtgcca gcgcgtgtgtc yggaaacttca gaggagacac aatcaagggc accaggaaatt	3325
20	tccctgtcc ccaagtggct tacagtcata ttggaaaca attaaacaac aatgcaaagc	3385
	aggttgtaat taaacactta atttgtgtat tctaactgtt aaggacaattt gaagtctqtg	3445
	gactcaggga tagatctt cccttcclgl cataccttc cctccagaac actgtgtct	3505
25	ctccagccag agaggatata aagggtttttt attttagaqt gaggcaatct ggactccct	3565
	ctctctagtt agtagttca ctttgaggcc aatcaatgaa cctcgctggg cctcagtttc	3625
	ctaatactaga aaataagggg ttggctccag atgatctaca aatttctctg gcccctgaggc	3685
30	catggggagg agggtggcgc agggtggc cctagagaag gaggagatct gaatcccttga	3745
	gtgaaggaag agggagctt ggggtggaga aaccatagat ttttttttt taagtggaa	3805
35	aatggagggg aaggggaaaa acaaggttct cagnagtcg gcagaggagt ggctttgca	3865
	ggttgactct ggaaagccgt tagaaattat ttaccagaga agcaacactc aaaaactca	3925
40	ctttggagac gacgaacctg gaggggcayy cagaactgac cagaatgaga ggacaggaca	3985
	ccaccagagg ggtcgagaca tgaggcagtg aggtgtggat gaaaccaggc gcgattgca	4045
	aagcagaaac cccaggctt tgcacgtatgt tgaaacactc agggccttggt gctctctgaa	4105
45	gcacccgtt ggtgcctggg gacttacgga gcccayggct ggaccagtgc ttggaggcc	4165
	agagcatggg gtcctgcca tgaacacagg tccaggact tgccttgcata tagatttctc	4225
	tccaggggag agaattccgtg gaggaagcag ggtgtcaggg gttggggaggg gggcagtgt	4285
50	glalaaccag ctgtttgtt cccctgggtg aacaggcagg aaaagcttc agagactoga	4345
	ggggaaaggt cagccatcag gagccaggta gagggggcgtc tatggcagac gggccactqa	4405
55	accccatctt gttcaggctt ctcccttccc lggatcttat gggtaaatga caggctttagg	4465
	gtgaagaagg aggcttgaga tgattctagt agagcctgtc atqactatgt agggggatgt	4525
60	ggggatggca ttgttcaqac ccaaggaggt gaaaccctcc ccaggatccc tcttttact	4585
	gtgaccacag gacaatgtac atgtggtcag caaatggatgtt cttgtatgtcc cgactctgtat	4645
	cgtttgcactt tggccagggtg acctccctc cctgggtctc aattttgacc tcatcaagat	4705
65	acagggttag actagttgtat ctgcatttgc atgacccatgaa attgatttgc tcccaatgtt	4765
	ccggaaatgtc ctggccccc tggccatggca aataaatgtca gggaccaact gatgaggc	4825
	gactttccca gccagccctg atacttctata gtagaggctt tgcatcagac cccacgaggc	4885

	caggagcaca gcaagggtcg tggagggct tagagctacc ccacitctag ctgatgaaac	4945
5	cctgcgtgc caccctactc ctccctacett tgcacagctc ttctctctt ggccactttc	5005
	aaaccacgcc tggcataact cagaatgcc ttctgcattc gtcagctga tgccagatgc	5065
	ctagtgtata ctggcgctg ttttgglgc tgtaagaagt caccatgtga tggtgtccct	5125
10	gctctcatgc atgtgtggag tcaccctggt gttagtctat ggataggggg tgctactga	5185
	gagagtcagg ggagatattc aaacagcaca agcataactg gtccccaga taggaggaac	5245
15	agagagaccc ctgaaatctt aacagcctt actgggtccct tgacaaggg ctggccggcc	5305
	agcgtcacag gagactatg tgacgctgtt cccaaaggcc acactttgcc tttccagatt	5365
	tgcagatgcc attatcttag gcctgtaca tgctggctg tctccctctg attctaagtg	5425
20	aacacccctc tgagggggac atctttctt ccaggcagat ttcaatttgc ccttgtacca	5485
	ttgcccctgt gtgcggcaca ctgcccagacg tccctgaact acttctttgc tcacatctgc	5545
25	ctgcccacca ctgcccattc gggcgctca ttctccctt cccagacccg caatgccttg	5605
	cccaacttcc aagcaccagt cccggctgtt ggccctccat cacatcttca aqaacaaccc	5665
	cacaaggctg atccgcttag ccaggcttcc cccaaacctgc tggggcattc tctagtctct	5725
30	atcgctttctt ccactcacac agcatttggt gtgcctcccc agacatctgc tttgtatctt	5785
	tttgtatttc taaaatcaattt gcccattcg tggggcccccc ttccaccatg agactatcag	5845
35	ttccttgaga gaagcgctca cctctgcctt ttcccttcc tcccacctgt ccctgtgcc	5905
	agccccactgt tggctttcag ttaatgatgt tgctgtgtg tgagacacac ccacatatgc	5965
	acatccacaa gccattcaca cacatggatc tgcacgttcc cacacccca catttgcaaa	6025
40	cacacgtcag ttgcctcag cacctgagaa ggcagccacc tctgtgagta ccattgcccac	6085
	aagtgacagt ctctgcattt atcgacgtcc tccctgtccc aatcccaggc aggcaaggct	6145
45	gcctctccac cagcccttcc accagccacc ctggcccttc ctggcccttc ctttgcacc	6205
	catgtcaaqq qggcaggaca gtggccatca cccctacccc aqggcaaaacg aggaaagcaa	6265
	agcctgaaac agtgagtggc atgacttacc aaggggctcc acagtttaca gccatttgac	6325
50	tttcttctga cllccaggga agggtgttt gcacctgcca ccgccttgcc tacatcccc	6385
	ctctctgtac ccagacttctt gggccccagt ggggatacag acagagraca gttgcattct	6445
55	gtgttccag gcctcttgg ccagcagcag ctggccagaa tggtccacgc acacaccc	6505
	tccctggctg tggctgagcc tcaagnatctt ggggagctgc ggltllyllycc acaccacacc	6565
	tccgtcacct caccctagg tgcacaaggc cagccctcaca ctccctttgc aagcgagcaq	6625
60	gggaaalccc ctctctgccc cacgagagga catgccacag tttgttccag aagccactcc	6685
	aaggccgagc gcccaggatg gtgtcaacgt gtcggaaact aacgggcaag agaggcttgg	6745
65	gtcaccaccc ccaggccagc ctcaggcccc tcagagggga aacttgcata ctgcacaaatg	6805
	ggagtccctgg gccaggcage calaagccgc tccclaycac cacaggcacc caaattcggc	6865
	agctcagagg tggaaaccacc caccctacagg ggtcagagag gagagacggg cagggacttq	6925

	ccacataggg ggttgaagtc ttgcttgtgg catctgaagg catgcaggcc tcgggtggtg	6985
	agaaaagaagg ggaatcatgg ggccaagggc agcagagccc cacgcagctt caagagttagg	7045
5	agatgttgc gcagagggat gaggctggga ggaaatccac cagcttagta gcagaacaca	7105
	gcagcacatg gtgtcaggaa gca tgaggaggg aagaggcctg agaggaaagg gggccctgcc	7165
10	tggagaccag cccccccct cagactgccc cccaccagcg cacacacaga cagacacaaa	7225
	cacactcagc ccaccctcg t gaccccccatt ggtggttcac ctggccca gcccagcatcc	7285
	cccaggccc agccctgat cactgtgtgg gttgtggca yyagggcgcc ttccctccaaa	7345
15	ggggccccc accctggccg agcgggtaga actggggcc tcggcagccac ccagctgggg	7405
	ccatccccggg caacacccag acaggcaccc aggctggaa ctgcctgaaa tctccatgag	7465
20	ccccaggcgg acctcctcaa gaagcaagtc ccagtcacg cgaacatcag ccatcccaag	7525
	caatgttcaa cacatcgta agaaaaccaca cagataagcc cccagattct accccaaact	7585
	ccagccaatt gttgcacta ccaaataattt gaaattatca ctggaaagcta aactggcagc	7645
25	ttgaagtctc ccatcctctt gatcccttcc ctgc cccac cccagctgca gaaggagga	7705
	ggaggcctag qggatgcact ggctggataa ttagaaccag cttccactct tcagccctcc	7765
30	tgggtccctct ctccagagga gacgcaggc acccctcacc accccgaaac tctgctggca	7825
	ccccctccct ctcaactctg ccaaagtcaa agtgc cccccc actacccqaa agcctctgga	7885
	gccccgacta tcagccactc agcaa atctc ctc cccggagt aattggatgt gatggAACGA	7945
35	gggctttca ggtctgttgtt ggaagtctctt ggtatgtttaatggcaca aaaaataaaaa	8005
	gtctcaqqcc tccccctac caccgcggg acatcattaa tctcggcaga aaggtatcag	8065
40	tgttttatac aatgtggccg gtccctgtac gagccagaca aattacagac ctgagccggg	8125
	ciclcgtgg ggcggattt aggctggagt tggatttttt ttttttctct tcagggccac	8185
	catcccccaa caagaagcag tggtgcctt acctaagagg agaggcagcg aggccaggga	8245
45	agagggtgtat cacctgggtt agatacagat caccacccly aggaacaggg ctgttcggag	8305
	ccggagccag aaccagtggc ccctgtcctt acagagaccc ccctgggaca agaggctgcc	8365
50	gcagctccaa cctcaaggc tcagcaaccc actccttgc tccctggccca cctgaacggg	8425
	ggcccatggg agaatggagc tgggggtgggg ggagctccc tccctcagct tccctacccca	8485
	cttcccaggc taagagggaa gtacatcagg atggaaatac atcaggacgg aaataactgag	8545
55	gaccaaggac cagtcagggtg ggaaacqclcc cttggctaca gcagggtcac ccaggaactg	8605
	ctgtcctcgg ctgttggaa gtcttaagag tggcctgtc taggtgcagt gactcacacc	8665
60	cgtaatccca gcaactttggg aggcccaggc aggaggatca ctggggccag gatgttgaga	8725
	tcagccctggca caacataggg agaccctgtc tctacagata aillttttt aattacccag	8785
	gtglqgtagc acacaactgt ggttccagct actcaggagg ctgaggcaag aggacccat	8845
65	gagcctggga ggtcaaggct gtaal gajag gtgatagtgc cactgcaccc cagcctggag	8905
	aatggcctcg ccgaaggaa tccagggtg tcactaggg tggggacctt gtclgclagg	8965
	gacccagact acaaagtct actgtggcc aaggccaccc tttgaaggac cagactgtga	9025

	ttaggttagt attacttgt a tatattaga aactcgccaa gggcaggggg taaaataagg	9085
5	cagctccgct gacccatgc ctggcagagg accagggtga tggctgccag tqttasaggg	9145
	gcctgggt ctgacccat ctaagagctt gcacttctgg cagggagcc gatccccat	9205
	tctaccaagt cctgcaaggg gcccaggta agctggatgt ggcagagcac ctctcagtct	9265
10	ctcagatgcc tgcttagccc taactggccc aatggagccg tccagagttt gtccagcacf	9325
	tgacaaacca ctcccaaaca gtggaaagcc aaccccacag atttgggaag cttggggtgg	9385
	agagagtgc acgtcaagacc tgccccagg cctttccaa acattcctgc catcatcaca	9445
15	gtccctggqa cttaagctcc gaggaggagc ccctccca gctgctgaag acaggggaca	9505
	tctttccccc ctagagccta ataggaggcc aggcacagag tggggccca gaaatcccc	9565
20	agtggcttta gataaggta tcaggctctc cccttggccc ccacatqgaa catgctgaa	9625
	atctaacatc tctcatcact aacatcttc ctatcagcac tgaccagtaa gtgagagtgc	9685
	caggactget gcatctggct atgcaggatg agaacagcac ccctacggc lglgcaggga	9745
25	tcaacctgta catcatatgt ggcagccctg tcctgagatg cccaggccctc tgggtccca	9805
	ccacaggcga ggtggctcat gagacaggca tcaqacagct tcctctctt ttctcccttc	9865
30	agccaaggcc tcccaccacc aaggcctctg aagtgcac actgcacca ccctcgccc	9925
	aggaagagcc tggccttagt gcatccagga tacagtggag tggcccaagc acagacagca	9985
35	gegctgcct ttccctgcct ggccgeggac tctctgcctc tcggggacag cagtgccata	10045
	caqgaagggg ctttccctcg agcggtgacc acagaagccg cagctctca qctctccagc	10105
	tctccccaca caccccaagg gcctgactca cccctctgc ttctagaac ttgcctctg	10165
40	tggcaggaaa caggggtgt aaggagtgc ggccaaagcca ggaaaggagg cagcaagagg	10225
	ctagccggca cggggactgg gtttcaggag ctggccaggg agcccgagaa ccacaggct	10285
	gtcgacttgg gactggaaag gaaccacgag tgccccagg cactgaaggc tcagaggctt	10345
45	ccgaacattt gaggggagcc agctctcgct ctggactcaq gggatttctt ccccgagct	10405
	tgccttgcct tctgtatggg ttcaattcc ttctctgccc tgcaggttcc ttctgcct	10465
50	acggtgggaa aggccccctc gtccctgttc ctgcacacca ccaattcacc cataagccg	10525
	tattaccctg gcatgtgcgc gtgcgcacac acacgtgcac ccacacatac acatgtgt	10585
	gcacacactc ttccctcaggc tcataataacc ataattcctt tcctccagc cagctggct	10645
55	ccctagggtgc agaggctcag gcacagcccc tctgaaagga cactcayglg acactgagcc	10705
	tccttaggaga acccgctgc aglcatgtaca cttccaccgg gagaggcttt ctcagcacct	10765
60	tggccaccc ttctgaaaaa gaaactcctgg aaccccttgc ttccctgccc ccctggccc	10825
	catcgatgc tgcagtcact gagtcgata ggatttgagg gaatctgcaa aggcccttcc	10885
	ccatccctgc cactacgaag cttcccttgc agtacccag cttgtgcctc agtcatccct	10945
65	tcgtgtact gggatggga acagccctcc cttttggga ttggltgtgag gattaagcga	11005
	attatgtgtc aagctttac agtagtacca acacataggt ggggtgttac tgctallatt	11065

60394576 .070000

gtttgcatta tlgltlaltgt tattttgggt ttcaagattaa aaattatcct ctccttgaga 11125
acatcttga aatgtcagca gcctctggga ggcagcaacg cttcaatggg cttgttatct 11185
5 taattgtat gaaaagggtg gtgggtgctg accgaggaga caggcaggtg ggagggtggg 11245
tggcccgta gtctcaagtg gggttcttag aycllccaac cccttgcctt ctctgtctg 11305
10 ggtatcagag gtgactggc tccattaatc cactgtgtat cgaccactgc ctagctctgc 11365
tccggacagg atctccccca ggactctct ctccaccctg ggaggttgct gcatttagca 11425
aaacaacaac aacaacaaaa agtaggatgc ccagtaaat tlyaalccca gataaacaac 11485
15 gaaaaaaaaa attcatgtaa ggatgtccca tgcagaatgt gggacacgct cacaccaaca 11545
agcgatttgc tattttatcg agattcaaat taaactactt ctcgcctttt cctggcaagc 11605
atctgcccag ggtgcctgac atccagggtt tggaaacaa gaacgtgggaa atatgcctct 11665
20 qaccacttcc ctcagctca ctcatgccc caggctgtgc ctttttgcc ctgccacaag 11725
gcctgttcac acttgcucca gaagatgcgc ctcctgtac acacttgtt 11785
25 tgaggaggaa cccctccgac ctgggcacccg tctgggttgc tgcaggccca gtcaccccg 11845
cagggtccag qggctgcccc tccacctctg ccatggggaa ggcacactgg tgcacccccc 11905
gtcctgggac agagacgagg aggtgtttgg ggttttaggg agaagattga aagaagagca 11965
30 cagatacttc caccaatctc agccccagca gcaggggtga gcaggtgtga aggaaagttag 12025
tgggtgaagc ctctgttctt ggggtcaggc gagcctgcag ggagaggc 12085
35 aagtgggcag cccgcacatgac tgggtttctt ttctgcccgg gagcgggagc gagaaggcag 12145
ttttqctaaa cagaggtgtt ggggaaggag attcagacat tctcaactgg ggttgggctc 12205
40 tcagttccca ttgcctctg gggagcagtc agctctgtg ctgttcagag gctctgtctaa 12265
catclggggg ctacctctgc aacaggagcc ggaagctccc ccaqccacta cgccccatgg 12325
ccatcgcacatc acctgtgc 12385
45 gctttcaagt ccattgttctt ttctcaactgc ctccctlggg tctgtttgtc cttccaggc 12444
tct tgg aaa atg gac aga ttt ctg aac cgt ttc cac ctg ggc gaa cct 12492
Ser Trp Lys Met Asp Arg Phe Leu Asn Arg Phe His Leu Gly Glu Pro
40 45 50
50 gaa gca agc acc cag ttc atg acc cag aac tac cag gac tca cca acc 12540
Glu Ala Ser Thr Gln Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro Thr
55 60 65
55 ctc cag gct ccc aga gaa agg gcc agc gag ccc aag cac aaa aac cag 12588
Leu Gln Ala Pro Arg Glu Arg Ala Ser Glu Pro Lys His Lys Asn Gln
70 75 80 85
60 gcg gag ata qct gcc tgc tag atagccggct ttgcctatccg ggcacatgtggc 12639
Ala Glu Ile Ala Ala Cys
90
cacactgcaccaccgacg atgtgggtat ggaacccctt ctggatacag aaccccttct 12699
65 tttccaaata aaaaaaaaaat catccaggc ttgggtgtt gtatccatgc ttgtattcc 12759
11 gaaaagca acaggaggaa attggccaat atccagcaat gggactgag gtttctgt 12819
aaatgcaagg qglqqgtgg ggagtggggg gaalagacgg acaggcagga tatagtgtac 12879

cctatgcctg aactgtcaact gagtcgtgt gagacacctac ggaaacacctct taccctcccc 12939
 5 aaagttctgc tcccttggtt gtgacacggg ggttaggaaga agtggggagag ggtgggttgt 12999
 ttatcaccac gggccttcc agctctgaca ttgcacaaacg ttcatteagc gactattgtt 13059
 gcatgtgggg tgcacccgtc ctccctccag gcactggaga tgcatgcctg cttcatcccc 13119
 10 atgcatgggt agatagggca gacacagcat ggtgctcacc tcataggagct tccagtccaa 13179
 caggggaagc aaatatcccc aacaaatgtat aaactgtgtat aagtatgtatg agggtgc当地 13239
 15 gggggcctga ttctgtacgg agagaaaaggg ttgaagactc aggcagagcc ctgtatgttta 13299
 acttcagggc tgaaaagtgg gctaggggtt accaggttag tggagacagg aatagagaac 13359
 attccaaagc ttctccgagc ccctgtggct ggcaggggca tgcatggcat cttccaggaa 13419
 20 atgaaaaggcc catgctgtqq ctgcgc当地 13479
 gggcaaaggc cagaccatgc caggcctttt ttgc当地 13539
 25 gaadgcaglg ggaaatcatt aaatgataag tagattggc caggcatggt ggctcggtcc 13599
 tataatcccc gtgcttggg aggccaaaggc cagaggatcg cttqaaqcca agagttcaag 13659
 accagcctgg gcagcatagg gagaccccca tccctacaaa aataagaata aaaaattagc 13719
 30 caggtatggt ggtggatgcc ttagtccca cctactcagg aggctggcat aggagaatca 13779
 cctgggccc当地 13839
 35 ccagecctgg caacccagca agacccatc tctaaaaat tcataaaata aaataaaata 13899
 aaaataqatt qqagagggcag atggcataat cagatgtatg ctttttagaaa gctgtgaata 13959
 acgccagggtg cggtggtctca cacctataat cccagcactt tggaaaggccaa aggtgggcaa 14019
 40 alcacclgag gtcaggagtt cgagatcagc ctgactaacg tggaaaccctg tctgtactaa 14079
 aaatac当地 14139
 45 tgaagcagga gaatcgctt aacccgggag gcagagggtt cagtgagccg agattgcacc 14199
 actgcactcc agctggc当地 14259
 aaaaaaaaaa gaaaagaaaag cggtaataa attcttaggt atctgcaagg aaggacttt 14319
 50 tgccatctaa cttcc 14334
 <210> 10
 <211> 131
 <212> DNA
 55 <213> Homo Sapiens
 <220>
 <221> misc-feature
 <222> (1)..(131)
 60 <223> exon 1
 <220>
 <221> CDS
 <222> (1)..(129)
 65 <400> 10
 atg ggt tgg aca atg agg ctg gtc aca gca gca ctg tta ctg ggt ctc
 Met Gly Trp Thr Met Arg Leu Val Thr Ala Leu Leu Leu Gly Leu
 1 5 10 15

66394576 . B7080Z

atg atg gtg gtc act gga gac gag gat gag aac agc ccg tgt gcc cat 96
 Met Met Val Val Thr Gly Asp Glu Asp Glu Asn Ser Pro Cys Ala His
 20 25 30
 5 gag gcc ctc ttg gac gag gac acc ctc ttt tgc ca 131
 Glu Ala Leu Leu Asp Glu Asp Thr Leu Phe Cys
 35 40
 10 <210> 11
 <211> 127
 <212> DNA
 <213> Homo Sapiens
 15 <220>
 <221> misc_feature
 <222> (1)..(127)
 <223> exon 2
 20 <220>
 <221> CDS
 <222> (2)..(127)
 25 <400> 11 49
 g qgc ctt gaa gtt ttc tac cca gag ttg ggg aac att ggc tgc aag gtt
 Gly Leu Glu Val Phe Tyr Pro Glu Leu Gly Asn Ile Gly Cys Lys Val
 1 5 10 15
 30 gtt cct gat tgt aac aac tac aga cag aag atc acc acc tcc tgg atg gag 97
 Val Pro Asp Cys Asn Asn Tyr Arg Gln Lys Ile Thr Ser Trp Met Glu
 20 25 30
 35 ccg ata gtc aag ttc ccg ggg gcc gtg gac 127
 Pro Ile Val Lys Phe Pro Gly Ala Val Asp
 35 40
 40 <210> 12
 <211> 99
 <212> DNA
 <213> Homo Sapiens
 <220>
 <221> misc_feature
 <222> (1)..(99)
 <223> exon 3
 45 <220>
 <221> CDS
 <222> (1)..(99)
 <400> 12
 ggc gca acc tat atc ctg gtg atg qtg gat cca gat gcc cct agc aga 48
 Gly Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp Ala Pro Ser Arg
 55 1 5 10 15
 gca gaa ccc aga cag Aga ttc tgg aga cat tgg ctg gta aca gat atc 96
 Ala Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu Val Thr Asp Ile
 60 20 25 30
 aag
 Lys 99
 65 <210> 13
 <211> 46
 <212> DNA
 <213> Homo Sapiens

60394576 .07000E

<220>
<221> misc_feature
<222> (1)..(46)
<223> exon 4
5
<220>
<221> misc_feature
<222> (16)..(16)
<223> predicted SNP site
10 a is replaced by g in seqID 17 (176..176)
and in seqID 18 (115..115)

<220>
<221> CDS
15 <222> (1)..(45)

<400> 13
ggc gcc gac ctg aag aaa ggg aag att cag ggc cag gag tta tca g 46
Gly Ala Asp Leu Lys Lys Gly Lys Ile Gln Gly Gln Glu Leu Ser
20 1 5 10 15

<210> 14
<211> 114
<212> DNA
25 <213> Homo Sapiens

<220>
<221> misc_feature
<222> (1)..(114)
30 <223> exon 5

<220>
<221> CDS
<222> (3)..(113)
35
<400> 14
cc tac cag gct ccc tcc cca ccq gca cac agt ggc ttc cat cgc tac 47
Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe His Arg Tyr
1 5 10 15
40 cag ttc ttt gtc tat ctt cag gaa gga aaa gtc atc tct ctc ctt ccc
Gln Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser Leu Leu Pro 95
20 25 30

45 aag gaa aac aaa act cga g 114
Lys Glu Asn Lys Thr Arg
35

<210> 15
50 <211> 269
<212> DNA
<213> Homo Sapiens

<220>
<221> misc_feature
<222> (1)..(269)
<223> exon 6 for seqID 3 and seqID 4
55

<220>
<221> CDS
<222> (3)..(155)

<400> 15
gc tct tgg aaa atg gac aga ttt ctg aac cgc ttc cac ctg ggc gaa 47
Ser Trp Lys Met Asp Arg Phe Leu Asn Arg Phe His Leu Gly Glu
65 1 5 10 15

cct gaa gca agc acc cag ttc atg acc cag aac tac cag gac tca cca 95
Pro Glu Ala Ser Thr Gln Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro

60394576 ..070802

	20	25	30	
5	acc ctc cag gct ccc aga gga agg gcc agc gag ccc aag cac aaa acc Thr Leu Gln Ala Pro Arg Gly Arg Ala Ser Glu Pro Lys His Lys Thr 35 40 45			143
10	agg cgg aga tag ctgcctgcta gatagccgc tttgccatcc gggcatgtgg Arg Arg Arg 50			195
15	ccacactgct caccacccgac gatgtggta tggaaaccccc tctggataca gaaccccttc tttccaaat taaa			255
20	<210> 16 <211> 269 <212> DNA <213> Homo Sapiens			269
25	<220> <221> misc_feature <222> (1)..(269) <223> exon 6 for seqID 1 and seqID 2			
30	<400> 16 gc tct tgg aaa atg gac aga ttt ctg aac cgt ttc cac ctg ggc gaa Ser Trp Lys Met Asp Arg Phe Leu Asn Arg Phe His Leu Gly Glu 1 5 10 15			47
35	cct gaa gca agc acc cag ttc atg acc cag aac tac cag gac tca cca Pro Glu Ala Ser Thr Gln Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro 20 25 30			95
40	acc ctc cag gct ccc aga gaa agg gcc agc gag ccc aag cac aaa aac Thr Leu Gln Ala Pro Arg Glu Arg Ala Ser Glu Pro Lys His Lys Asn 35 40 45			143
45	cag gcg gag ata gct gcc tgc tagatagccg gctttgccat ccggcatgt Gln Ala Glu Ile Ala Ala Cys 50			194
50	ggccacactg cccaccacccg acgatgtggg tatggaaaccc cctctggata cagaacccct tctttccaa ataaa			254
55	<210> 17 <211> 624 <212> DNA <213> Homo Sapiens			269
60	<220> <221> misc_feature <222> (1)..(624) <223> public sequence BM984582 from dbEST (1..624 c)			
65	<220> <221> CDS <222> (2) ..(487) <223> seqID 1 (67..227) seqID 2 (45..205)			
	<400> 17 c tac aga cag aag atc acc tcc tgg atg qaq ccg ata gtc aag ttc ccg			49

60394576 . G703DE

Tyr Arg Gln Lys Ile Thr Ser Trp Met Glu Pro Ile Val Lys Phe Pro
1 5 10 15

5 ggg gcc gtg gac ggc gca acc tat atc ctg gtg atg gtg gat cca gat 97
Gly Ala Val Asp Gly Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp
20 25 30

10 gcc cct agc aga gca gaa ccc aga cag aga ttc tgg aga cat tgg ctg 145
Ala Pro Ser Arg Ala Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu
35 40 45

15 gta aca gat atc aag ggc gcc gac ctg aag gaa ggg aag att cag ggc 193
Val Thr Asp Ile Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile Gln Gly
50 55 60

20 cag gag tta tca gcc tac cag gct ccc tcc cca ccg gca cac agt ggc 241
Gln Glu Leu Ser Ala Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly
65 70 75 80

25 ttc cat cgc tac cag ttc ttt gtc tat ctt cag gaa gga aaa gtc atc 289
Phe His Arg Tyr Gln Phe Val Tyr Leu Gin Glu Gly Lys Val Ile
85 90 95

30 tct ctc ctt ccc aag gaa aac aaa act cga ggc tct Lgg aaa atg gac 337
Ser Leu Pro Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys Met Asp
100 105 110

35 aga ttt ttg aac cgt ttc cac ctg ggc gaa cct gaa gca agc acc cag 385
Arg Phe Leu Asn Arg Phe His Leu Gly Glu Pro Glu Ala Ser Thr Gln
115 120 125

40 ttc atg acc caq aac tac cag gac tca cca acc ctc cag gct ccc aga 433
Phe Met Thr Gln Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala Pro Arg
130 135 140

45 gaa agg gcc agc gag ccc aag cac aaa aac cag qcq gag ata gct gcc 481
Glu Arg Ala Ser Glu Pro Lys His Lys Asn Gln Ala Glu Ile Ala Ala
145 150 155 160

50 tgc tag atagccgct ttgccatccg ggcatgtggc cacactgcc accaccgacg 537
Cys

55 atgtgggtat ggaaccccct ctggatacag aacccttct tttccaaata aaaaaaaaaat 597
catccaggaa aaaaaaaaaa aaaaaaaaa 624

55 <210> 18
<211> 547
<212> DNA
<213> Homo Sapiens

60 <220>
<221> misc_feature
<222> (1)..(547)
<223> public sequence AI857902 from dbEST (1..547 c)

65 <220>
<221> CDS
<222> (1)..(414)
<223> seqID 3 (87..223)
seqID 4 (65..223)

65 <400> 18 48
ggc qca acc tat atc ttg gtg atg glg yat cca gat gcc cct agc aga
Gly Ala Thr Tyr Ile Leu Val Met Val Asp Pro Asp Ala Pro Ser Arg
1 5 10 15

50394576 . 070802

	gca gaa ccc aga cag aga ttc tgg aga cat tgg ctg gta aca gat atc Ala Glu Pro Arg Gln Arg Phe Trp Arg His Trp Leu Val Thr Asp Ile 20 25 30	96
5	aag ggc gcc gac ctg aag gaa ggg aag att cag ggc cag gag tta tca Lys Gly Ala Asp Leu Lys Glu Gly Lys Ile Gln Gly Gln Glu Leu Ser 35 40 45	144
10	gcc tac cag gct ccc tcc cca ccg gca cac agt ggc ttc cat cgc tac Ala Tyr Gln Ala Pro Ser Pro Pro Ala His Ser Gly Phe His Arg Tyr 50 55 60	192
15	cag ttc ttt gtc tat ctt cag gaa gga aaa gtc atc tct ctc ctt ccc Gln Phe Val Tyr Leu Gln Glu Gly Lys Val Ile Ser Leu Leu Pro 65 70 75 80	240
	aag gaa aac aaa act cga ggc tct tgg aaa atg gac aga ttt ctg aac Lys Glu Asn Lys Thr Arg Gly Ser Trp Lys Met Asp Arg Phe Leu Asn 85 90 95	288
20	cgc ttc cac ctg ggc gaa cct gaa gca agc acc cag ttc atg acc cag Arg Phe His Leu Gly Glu Pro Glu Ala Ser Thr Gln Phe Met Thr Gln 100 105 110	336
25	aac tac cag gac tca cca acc ctc cag gct ccc aga gga agg gcc agc Asn Tyr Gln Asp Ser Pro Thr Leu Gln Ala Pro Arg Gly Arg Ala Ser 115 120 125	384
30	gag ccc aag cac aaa acc agg cgg aga tag ctgcctgcta gatagccggc Glu Pro Lys His Lys Thr Arg Arg Arg 130 135	434
	tttgccatcc gggcatgtgg ccacactgct caccaccgac gatgtggta tggaaccccc	494
35	tttggataca gaaccccttc tttccaaat taaaaaaaaa aatcatccag gaa	547

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.